B. V. GNEDENCO şi A. I. HINCIN

INTRODUCERE ELEMENTARĂ IN CALCULUL PROBABILITĂŢILOR

TRADUCERE DIN LIMBA RUSĂ

EDITURA TEHNICĂ

Scanned with OKEN Scanner

B. V. GNEDENCO ŞI A. I. HINCIN

INTRODUCERE ELEMENTARĂ ÎN CALCULUL PROBABILITĂŢILOR

TRADUCERE DIN LIMBA RUSĂ

EDITURA TEHNICĂ
1953

Scanned with OKEN Scanner

Б. В. ГНЕДЕНКО и А. И. ХИНЧИН

ЭЛЕМЕНТАРНОЕ ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

ГОСТЕХТЕОРИЗДАТ
МОСКВА 1952 ЛЕНИНГРАД

Scanned with OKEN Scanner

PREFAȚĂ LA PRIMA EDIȚIE

Cunoașterea bazelor teoretice ale unui discipline matematice dă totdeauna posibilitatea ca rezultatele ei să fie aplicate în practică în mod mai conștient și mai activ. In domeniul calculului probabilităților, însă, lucrurile se prezintă în așa fel că un mare număr de cadre de conducere (și câteodată simpli muncitori) din armată, industrie, agricultură, economie, etc., cu cunoștințe matematice relativ restrânse, trebue să recurgă la diferite aplicații practice ale acestei discipline.

Broșura noastră are drept scop să facă cunoscute acestor cadre, în forma cea mai accesibilă cu putință, noțiunile elementare ale calculului probabilităților și metodele folosite în calculul probabilităților. Broșura este total accesibilă absolvenților școlilor medii de zece ani și aproape în întregime accesibilă absolvenților școlilor de șapte ani. In mai toate capitolele, expunerea se bazează pe exemple practice concrete, dar în alegerea acestor exemple nu ne-am condus după actualitatea lor practică ci, în primul rând, după caracterul lor ilustrativ, necesar însușirii problemelor teoretice corespunzătoare.

Moscova, 7 Ianuarie 1945.

Scanned with OKEN Scanner

PREFAȚĂ LA EDIȚIA A DOUA

Noua ediție a cărții oferită cititorului diferă numai într'o măsură neînsemnată de prima ediție: am adăugat numai încheierea și am introdus un mic număr de observații explicative în text.

Chiev—Moscova, 7 Noembrie 1949.

PREFAȚĂ LA EDIȚIA A TREIA

Ediția a treia este aproape identică cu a doua: s'au făcut doar câteva îndreptări neînsemnate.

17 Aprilie 1952.

Scanned with OKEN Scanner

PARTEA ÎNTÂIA

PROBABILITĂȚI

CAPITOLUL I

PROBABILITĂȚILE EVENIMENTELOR

§ 1. Noțiunea de probabilitate

Când spunem despre un trăgător oarecare că nimerește ținta în proporție de 92% la o tragere în condiții date, aceasta înseamnă că, din 100 de focuri trase, executate în condiții determinate (aceeași țintă la aceeași distanță, aceeași armă, etc.), *în medie* cam 92 de focuri sunt reușite (și deci cam 8 nereușite). Desigur că nu la fiecare sută de focuri vom avea 92 focuri reușite; acestea vor fi câteodată în număr de 91 sau 90, câteodată 93 sau 94; uneori se poate întâmpla chiar ca numărul lor să fie mult mai mic sau mult mai mare decât 92; dar *în medie,* dacă se repetă tragerea de multe ori în aceleași condiții, acest procent de nimeriri la țintă va rămâne invariabil, atâta timp cât nu vor interveni unele modificări esențiale (astfel, trăgătorul nostru își poate ridica măestria, iar procentul mediu de focuri reușite se poate urca până la 95% și mai mult). Experiența arată că, pentru un astfel de trăgător, în majoritatea cazurilor, la 100 focuri trase, numărul focurilor reușite va fi aproximativ 92; cazul în care, la 100 de focuri trase, acest număr va fi, de exemplu, mai mic decât 88 sau mai mare decât 96 se va putea totuși întâlni, dar relativ rar. Cifra de 92%, care servește drept indice al măestriei trăgătorului nostru, este deobicei foarte *stabilă,* adică procentul nimeririlor la țintă, în cele mai multe cazuri de tragere (în aceleași condiții) va fi același pentru un trăgător, abătându-se mult dela valoarea lui medie numai în cazuri excepționale.

Să mai considerăm încă un exemplu. La o întreprindere oarecare s'a remarcat că, în condiții date, în medie 1,6% din obiectele fabricate sunt necorespunzătoare standardului și merg la rebut. Aceasta înseamnă că într'o partidă dată de mărfuri, care nu au fost încă sortate, formată să zicem din 1 000 piese fabricate, vor fi,

Scanned with OKEN Scanner

de exemplu, 16 piese nesatisfăcătoare. Desigur, câteodată numărul rebuturilor va fi ceva mai mare, altădată ceva mai mic, dar în medie acest număr va fi apropiat de 16 și, în cele mai multe partide de mărfuri de câte 1 000 piese, numărul celor date la rebut va fi apropiat de 16. Se înțelege că și aici presupunem condițiile de fabricație (organizarea procesului tehnologic, utilajul, materia primă, calificarea muncitorilor, etc.) invariabile.

Este clar că se pot cita oricâte exemple de acest fel. In toate aceste exemple, se vede că, la operații *de masă, omogene* (focuri trase repetat, fabricarea în serie a produselor, etc.), procentul unuia sau altuia dintre evenimentele care ne interesează (ținte nimerite, produse necorespunzătoare standardului, etc.), va fi aproape totdeauna același, în condiții date, abătându-se într'o măsură mai însemnată dela o anumită valoare medie numai în cazuri rare. Deaceea, se poate spune că această valoare medie este indicele caracteristic al operației de masă date (în condiții date și stabilite cu precizie). Procentul de ținte nimerite exprimă gradul de măestrie a trăgătorului, procentul rebuturilor arată calitatea producției. Deaceea, se înțelege dela sine că cunoașterea unui astfel de indice este foarte importantă în cele mai variate domenii: știință militară, tehnică, economie, fizică, chimie, etc.; ea ne permite să apreciem fenomene de masă deja întâmplate și chiar să prevedem rezultatul viitor al unei operații de masă sau alteia.

Dacă trăgătorul, în condițiile de tragere date, nimerește ținta în medie de 92 ori la 100 de focuri trase, spunem că, pentru acest trăgător și în aceste condiții, *probabilitatea nimeririi* țintei este de 92% (sau 92/100, sau 0,92). Dacă, în condiții date, la 1 000 piese fabricate de o întreprindere oarecare revin în medie 16 piese de rebut, spunem că *probabilitatea de a produce rebuturi* este pentru întreprinderea dată de 0,016 sau 1,6%.

Ce înțelegem, în general, prin probabilitatea unui eveniment într'o operație de masă dată? Acum este ușor de răspuns la această întrebare. O operație de masă constă totdeauna din repetarea unui număr mare de operații individuale, asemănătoare între ele (tragerea — din focuri trase separat, producția de masă — din fabricarea de produse separate, etc.). Pe noi ne interesează un rezultat determinat al unei singure operații individuale (nimerire la țintă la un singur foc tras, necorespunderea cu standardul a unui singur produs separat, etc.) și înainte de toate — numărul unor astfel de rezultate într'o operație de masă sau alta (câte focuri trase vor nimeri la țintă, câte produse vor fi rebuturi, etc.). Procentul (sau în general fracția) unor astfel de rezultate

Scanned with OKEN Scanner

„reușite“ [1] din operația de masă dată îl denumim *probabilitatea* acestui rezultat — important pentru noi. In acest caz, trebue să se aibă totdeauna în vedere că problema determinării probabilității unui eveniment (rezultat) sau altuia are sens numai dacă sunt determinate cu precizie condițiile în care se desfășoară operația noastră de masă. Orice modificare esențială a acestor condiții atrage după sine, de regulă, modificarea probabilității care ne interesează.

Dacă operația de masă este de așa natură încât evenimentul A (de exemplu, nimerire la țintă) are loc în medie de a ori în b operații individuale (focuri trase), probabilitatea evenimentului A, în condiții date, este $\frac{a}{b}$ (sau $\frac{100\,a}{b}$ %). Deaceea putem spune că probabilitatea rezultatului „reușit“ al unei operații individuale este *raportul dintre numărul mediu de rezultate „reușite“ observat și numărul total al operațiilor individuale* care alcătuesc operația de masă dată. Se înțelege dela sine că, dacă probabilitatea unui eveniment oarecare este egală cu $\frac{a}{b}$, la fiecare serie dată de b operații individuale, acest eveniment va putea avea loc și mai mult decât de a ori și mai puțin decât de a ori; numai în medie el are loc de a ori; și, în cele mai multe din astfel de serii de câte b operații, evenimentul A va apare de un număr de ori apropiat de a, *mai ales dacă b este un număr mare.*

Exemplul 1. Intr'un oraș oarecare, în cursul primului trimestru s'au născut:

în Ianuarie — 145 băieți și 135 fete,
în Februarie — 142 băieți și 136 fete,
în Martie — 152 băieți și 140 fete.

Care este probabilitatea nașterii unui băiat? Procentul băieților nou născuți este:

$$\text{în Ianuarie: } \frac{145}{280} \approx 0{,}518 = 51{,}8\,\%,$$

$$\text{în Februarie: } \frac{142}{278} \approx 0{,}511 = 51{,}1\,\%,$$

$$\text{în Martie: } \frac{152}{292} \approx 0{,}520 = 52{,}0\,\%.$$

Vedem că media aritmetică a procentelor pe diferite luni este apropiată de numărul $0{,}516 = 51{,}6\,\%$; probabilitatea căutată este,

[1] In al doilea exemplu ar trebui spus mai curând „nereușite“. Totuși, se obișnuește în teoria probabilităților a denumi „reușite“ acele rezultate care duc la înfăptuirea evenimentului care ne interesează în problemă.

în condiţiile date, aproximativ 0,516 sau 51,6%. Această cifră este bine cunoscută în demografie — ştiinţa care studiază dinamica populaţiei; se constată că procentul naşterilor de băieţi, în condiţii obişnuite, nu va diferi simţitor de această cifră, în diferite perioade.

Exemplul 2. La începutul secolului trecut a fost descoperit un fenomen remarcabil denumit (după numele botanistului Brown care l-a descoperit) mişcarea browniană. Acest fenomen constă în faptul că particulele foarte mici în suspensie din lichid [1] se află într'o mişcare dezordonată, care se produce fără nicio cauză vizibilă.

Multă vreme nu s'a putut lămuri cauza acestei mişcări, care părea arbitrară, până când teoria cinetică a gazelor a dat o explicaţie simplă şi completă: mişcarea particulelor în suspensie, în lichid, este rezultatul ciocnirii moleculelor lichidului de aceste particule. Teoria cinetică a gazelor dă posibilitatea de a se calcula probabilitatea ca într'un volum dat de lichid să nu existe nicio particulă în suspensie, probabilitatea ca în lichid să existe o asemenea particulă, două, trei, etc. In scopul verificării rezultatelor acestei teorii, s'au efectuat o serie de experienţe.

Vom da rezultatele a 518 observaţii făcute de fizicianul suedez Swedeberg asupra particulelor celor mai mici posibile de aur, suspendate în apă. S'a constatat că în porţiunea supusă observaţiei nu s'a observat nicio particulă de 112 ori, o particulă s'a observat de 168 ori, 2 particule — de 130 ori, 3 particule — de 69 ori, 4 particule — de 32 ori, 5 particule — de 5 ori, 6 particule — odată, 7 particule — odată.

Procentul unui număr sau altuia de particule observate este:

0 particule: $\frac{112}{518} \approx 0{,}216$;

1 particulă: $\frac{168}{518} \approx 0{,}325$;

2 particule: $\frac{130}{518} \approx 0{,}251$;

3 particule: $\frac{69}{518} \approx 0{,}133$;

4 particule: $\frac{32}{518} \approx 0{,}062$;

5 particule: $\frac{5}{518} \approx 0{,}010$;

6 particule: $\frac{1}{518} \approx 0{,}002$.

7 particule: $\frac{1}{518} \approx 0{,}002$;

Rezultatele observaţiilor s'au dovedit a coincide cu probabilităţile prevăzute teoretic.

---

[1] Deci care se află de echilibru indiferent.

Scanned with OKEN Scanner

§ 2. Evenimente imposibile și evenimente sigure

Este evident că probabilitatea unui eveniment este totdeauna un număr pozitiv sau zero. Ea nu poate fi mai mare decât unitatea pentrucă în fracția prin care se exprimă probabilitatea, numărătorul nu poate fi mai mare decât numitorul (numărul operațiilor „reușite" nu poate fi mai mare decât numărul tuturor operațiilor efectuate).

Să convenim a nota prin $P(A)$ probabilitatea evenimentului A. Oricare ar fi acest eveniment,

$$0 \leqslant P(A) \leqslant 1.$$

Cu cât $P(A)$ este mai mare, cu atât mai des are loc evenimentul A. De exemplu, cu cât la un trăgător probabilitatea nimeririi la țintă este mai mare, cu atât mai multe sunt focurile reușite trase de el, cu atât mai mare este măestria sa. Dacă probabilitatea evenimentului este foarte mică, el are loc rar; dacă $P(A)=0$, evenimentul A sau nu are loc niciodată, sau are loc extrem de rar, așa că practic îl putem socoti ca imposibil. Dacă, dimpotrivă, $P(A)$ este aproape de 1, atunci în fracția prin care se exprimă această probabilitate, numărătorul are o valoare apropiată de cea a numitorului, adică majoritatea covârșitoare a operațiilor o constitue cele reușite; un astfel de eveniment are loc în cea mai mare parte a cazurilor; dacă $P(A)=1$, evenimentul are loc totdeauna, așa că practic îl putem socoti, cum se spune, „sigur", adică putem socoti că va avea loc cu siguranță. Dacă $P(A)=\frac{1}{2}$, evenimentul A survine cam la jumătate din numărul cazurilor; aceasta înseamnă că se constată tot atâtea operații „reușite" cât și „nereușite". Dacă $P(A)>\frac{1}{2}$ evenimentul A are loc de mai multe ori decât nu are loc; când $P(A)<\frac{1}{2}$, fenomenul este invers.

Cât de mică trebue să fie probabilitatea unui eveniment pentru ca să-l socotim imposibil din punct de vedere practic? Nu se poate da un răspuns general la această întrebare, întrucât totul depinde de cât de important este evenimentul despre care este vorba. Astfel, 0,01 este un număr mic. Dacă avem un grup de proiectile și probabilitatea ca proiectilul să nu explodeze la cădere este 0,01, aceasta înseamnă că aproximativ 1% din focuri vor rămâne fără rezultat. Cu acest rezultat ne putem considera satisfăcuți. Dar dacă avem o parașută și 0,01% este probabilitatea ca la lansare ea să nu se deschidă, acest rezultat, desigur, nu ne poate satisface căci ar însemna că la fiecare sută de lansări cu parașuta

Scanned with OKEN Scanner

va muri fără rost un parașutist. Aceste exemple arată că, în fiecare problemă, trebue să stabilim dinainte, pe baza unor considerente practice, cât de mică trebue să fie probabilitatea unui eveniment pentru ca, fără știrbire pentru cauză, să putem neglija posibilitatea realizării lui.

§ 3. Problemă

Problemă. Un trăgător nimerește ținta în 80% din cazuri, iar altul (în aceleași condiții de tragere) în 70%. Să se găsească probabilitatea nimeririi țintei dacă amândoi trăgătorii trag simultan în ea. Ținta se consideră atinsă dacă cel puțin unul din cele două gloanțe nimerește în ea.

Metoda I-a de rezolvare. Să presupunem că se trag 100 de focuri duble. Aproximativ 80 din cele trase de primul trăgător vor nimeri ținta. Rămân cam 20 de focuri pe care acest trăgător le va rata. Deoarece al doilea trăgător nimerește ținta cam de 70 ori la 100 de focuri trase și deci de 7 ori la 10 focuri trase, ne putem aștepta ca în cazul celor 20 de focuri trase în care primul va rata ținta, al doilea dintre ei va nimeri ținta de aproximativ 14 ori. Astfel, la 100 de focuri trase, ținta va fi atinsă de aproximativ $80+14=94$ ori. Probabilitatea atingerii țintei la o tragere simultană a celor doi trăgători este deci egală cu 94%, sau 0,94.

Metoda II-a de rezolvare. Să presupunem din nou că se trag 100 de focuri duble. Am văzut mai sus că, în aceste condiții, primul trăgător va rata ținta cam de 20 de ori. Deoarece al doilea trăgător greșește ținta cam de 30 ori la 100 de focuri trase și deci de trei ori la 10 focuri trase, ne putem aștepta ca între cele 20 de focuri ratate de primul trăgător să fie și 6 focuri ratate de către al doilea trăgător. La fiecare dintre aceste 6 focuri ținta rămâne neatinsă, iar în cazul celorlalte 94 focuri trase, cel puțin unul din trăgători va trage cu succes și deci ținta va fi atinsă. Ajungem din nou la concluzia că, la o tragere dublă, ținta va fi atinsă în cam 94 cazuri dintr'o sută, adică probabilitatea atingerii țintei este 94% sau 0,94.

Problema pe care am studiat-o este foarte simplă. Totuși, ea ne conduce la un rezultat foarte important: există cazuri când este util să putem deduce din probabilitățile unor evenimente, probabilitățile altor evenimente mai complexe. Asemenea cazuri sunt într'adevăr foarte numeroase nu numai în știința militară ci și în orice activitate practică, în care avem a face cu fenomene de masă.

Ar fi desigur foarte incomod ca pentru fiecare problemă nouă care ni se pune să căutăm o metodă a ei proprie de rezolvare.

Scanned with OKEN Scanner

Știința caută întotdeauna să creeze reguli generale, a căror cunoaștere să ne permită să rezolvăm mecanic, sau aproape mecanic, diferite probleme separate, care se aseamnă între ele.

In domeniul fenomenelor de masă, știința care se ocupă cu stabilirea unor astfel de reguli generale se numește *calculul probabilităților*. In această carte vor fi date primele elemente fundamentale ale acestei științe.

Calculul probabilităților este una din ramurile matematicei, ca și aritmetica sau geometria. Calea lui este deaceea calea raționamentului riguros, iar instrumentele lui sunt formulele, tabelele, desenul și altele.

CAPITOLUL II

REGULA DE ADUNARE A PROBABILITĂȚILOR

§ 4. Deducerea regulii de adunare a probabilităților

Regula generală cea mai simplă și cea mai importantă folosită în calculul probabilităților este *regula adunării*, pe care o vom studia acum.

La o tragere într'un panou, arătată în fig. 1, pentru fiecare trăgător și la o distanță dată, există o anumită probabilitate de a nimeri în fiecare dintre regiunile *1, 2, 3, 4, 5, 6*. Să presupunem că pentru un trăgător oarecare probabilitatea de a nimeri în regiunea *1* este de 0,24, iar probabilitatea de a nimeri în regiunea *2* este de 0,17. După cum știm, aceasta înseamnă că dintr'o sută de gloanțe trase de către acest trăgător, 24 gloanțe nimeresc (în medie) în regiunea *1* și 17 gloanțe în regiunea *2*.

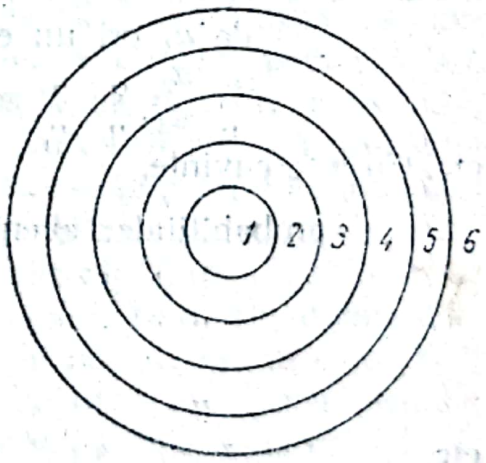

Fig. 1.

Să presupunem că, la un concurs de tir, focul tras se consideră „excelent" dacă glonțul a nimerit în regiunea 1 și „bun" dacă a nimerit în regiunea *2*. Care este probabilitatea ca focul tras de trăgătorul nostru să fie bun sau excelent?

Este ușor de răspuns la această întrebare. La fiecare 100 de gloanțe trase de către trăgător, cam 24 cad în regiunea *1* și cam

17 în regiunea 2. Deci, la fiecare 100 de gloanțe, vor fi cam 24+17=41 gloanțe care vor nimeri fie în regiunea 1 fie în regiunea 2. Probabilitatea căutată este deci egală cu 0,41=0,24+0,17. *Probabilitatea ca focul tras să fie excelent sau bun este deci egală cu suma probabilităților focurilor trase excelent și a celor trase bine.*

Să examinăm un alt caz. Un călător așteaptă tramvaiul Nr. 26 sau Nr. 16, la o stație la care opresc tramvaiele a patru linii: Nr. 16, Nr. 22, Nr. 26, Nr. 31. Presupunând că tramvaiele de pe toate aceste linii sosesc în medie cu aceeași frecvență, să se găsească probabilitatea ca primul din tramvaiele sosite în stație să fie unul din cele necesare călătorului.

Este limpede că probabilitatea ca primul din tramvaiele sosite în stație să fie Nr. 16 este egală cu $\frac{1}{4}$. Aceeași probabilitate există și pentru ca primul tramvai sosit să fie Nr. 26. Probabilitatea căutată este deci egală, după cum se vede, cu $\frac{1}{2}$. Dar

$$\frac{1}{2}=\frac{1}{4}+\frac{1}{4};$$

putem deci spune că probabilitatea ca primul din tramvaiele sosite să fie Nr. 16 sau Nr. 26 este egală cu suma probabilităților sosirii tramvaiului Nr. 16 și a tramvaiului Nr. 26.

Putem face acum un raționament general.

In efectuarea unei operații de masă oarecare, s'a stabilit că în fiecare serie de b operații individuale se observă în medie

de a_1 ori un eveniment oarecare A_1,
„ a_2 „ „ „ „ A_2,
„ a_3 „ „ „ „ A_3,

etc. Cu alte cuvinte,

probabilitatea evenimentului A_1 este egală cu $\frac{a_1}{b}$,
„ „ A_2 „ „ „ $\frac{a_2}{b}$,
„ „ A_3 „ „ „ $\frac{a_3}{b}$

etc.

Care este probabilitatea ca într'o operație individuală oarecare să aibă loc unul din evenimentele A_1, A_2, A_3 (indiferent care)?

Evenimentul care ne interesează poate fi denumit A_1 sau A_2 sau A_3 sau ...[1]. Într'o serie de b operații, acest eveniment se

[1] Puncte, puncte aici și în cazuri asemănătoare înseamnă „și așa mai departe".

Scanned with OKEN Scanner

produce de $a_1+a_2+a_3+\ldots$ ori; deci probabilitatea căutată este egală cu

$$\frac{a_1+a_2+a_3+\ldots}{b}=\frac{a_1}{b}+\frac{a_2}{b}+\frac{a_3}{b}+\ldots;$$

aceasta se poate exprima prin următoarea formulă:

$$P(A_1, \text{ sau } A_2, \text{ sau } A_3, \text{ sau } \ldots)=P(A_1)+P(A_2)+P(A_3)+\ldots$$

Atât în exemplele noastre cât și în raționamentul general, am presupus tot timpul că oricare două din evenimentele examinate (de exemplu A_1 și A_2) *sunt incompatibile între ele*, adică nu pot avea loc în aceeași operație individuală. De exemplu, tramvaiul sosit în stație nu poate avea în același timp un număr necesar nouă și un număr de care nu avem nevoie — el sau satisface nevoile călătorului, sau nu le satisface.

Această ipoteză asupra incompatibilității reciproce a evenimentelor distincte care ne interesează este foarte importantă; fără de ea, regula adunării nu mai este adevărată și aplicarea ei ar duce la greșeli serioase. Să considerăm, de exemplu, problema rezolvată la sfârșitul paragrafului precedent (pag. 10). Am căutat, în acea problemă, probabilitatea ca la o tragere dublă ținta să fie nimerită fie de primul, fie de cel de al doilea trăgător, probabilitatea de a nimeri ținta fiind pentru primul trăgător 0,8 și pentru al doilea 0,7. Dacă am încerca să aplicăm în rezolvarea acestei probleme regula adunării, ar rezulta că probabilitatea căutată este egală cu $0,8+0,7=1,5$ ceeace este absurd, deoarece știm că probabilitatea unui eveniment nu poate să fie mai mare ca unitatea. Am ajuns la acest rezultat fals și absurd, pentrucă am aplicat regula adunării într'un caz unde ea nu este aplicabilă; cele două evenimente despre care este vorba în această problemă sunt *compatibile* între ele, pentrucă este foarte posibil ca amândoi trăgătorii să atingă ținta la același foc dublu tras. O mare parte din greșelile pe care le fac începătorii în calculul probabilităților se datoresc tocmai utilizării greșite a regulii adunării; trebue deci să evităm cu grijă această greșală și, în fiecare caz în care vrem să aplicăm regula adunării, să verificăm dacă toate evenimentele care intervin în acea problemă sunt incompatibile două câte două.

Putem da acum o formulare generală a regulii de adunare.

Regula de adunare. *Probabilitatea ca într'o operație oarecare să se producă un eveniment oarecare dintre evenimentele* $A_1, A_2, \ldots, A_n$ *(indiferent care din ele) este egală cu suma probabilităților acestor evenimente, dacă evenimentele sunt incompatibile două câte două.*

Scanned with OKEN Scanner

§ 5. Sistem complect de evenimente

In cei 20 ani de durată a celui de al treilea Imprumut de Stat pentru refacerea și desvoltarea economiei naționale, a treia parte din obligații ies câștigătoare, iar celelalte două treimi ies la tragere la sorți și se achită după valoarea lor nominală. Cu alte cuvinte, fiecare obligație din acest împrumut are o posibilitate de câștig egală cu $\frac{1}{3}$ și o probabilitate de ieșire la sorți egală cu $\frac{2}{3}$. Câștigul și ieșirea la sorți sunt *evenimente contrarii,* adică sunt două evenimente, din care unul și numai unul are loc în mod necesar pentru fiecare obligație. Suma probabilităților lor

$$\frac{1}{3}+\frac{2}{3}=1$$

și acest lucru nu este întâmplător. In general, dacă A_1 și A_2 sunt două evenimente contrarii și dacă într'o serie de b operații evenimentul A_1 se produce de a_1 ori, iar evenimentul A_2 de a_2 ori, atunci, evident, $a_1+a_2=b$. Dar

$$P(A_1)=\frac{a_1}{b},\quad P(A_2)=\frac{a_2}{b},$$

astfel încât

$$P(A_1)+P(A_2)=\frac{a_1}{b}+\frac{a_2}{b}=\frac{a_1+a_2}{b}=1,$$

Se poate obține același rezultat și din regula de adunare; deoarece evenimentele contrarii sunt incompatibile între ele, rezultă că:

$$P(A_1)+P(A_2)=P(A_1,\text{ sau }A_2);$$

dar evenimentul «A_1 sau A_2» este un eveniment sigur deoarece din definiția evenimentelor contrarii rezultă că el trebue să se producă cu siguranță; deaceea probabilitatea lui este egală cu unitatea și obținem din nou

$$P(A_1)+P(A_2)=1.$$

Suma probabilităților a două evenimente contrarii este egală cu unitatea.

Această regulă admite o generalizare foarte importantă care se poate demonstra în același mod. Fie n (un număr oarecare) de evenimente $A_1, A_2, \ldots, A_n$, astfel încât în fiecare operație individuală trebue să aibă loc unul dintre aceste evenimente și numai unul; convenim să numim un astfel de grup de evenimente *sistem*

Scanned with OKEN Scanner

complect de evenimente; în particular, fiecare pereche de evenimente contrarii formează, în mod evident, un sistem complect.

Suma probabilităților unor evenimente care formează un sistem complect este egală cu unitatea.

În adevăr, conform definiției sistemului complect, oricare două din evenimentele acestui sistem sunt incompatibile între ele, și deci din regula adunării rezultă:

$$P(A_1)+P(A_2)+\ldots+P(A_n)=P(A_1, \text{ sau } A_2, \text{ sau } \ldots, \text{ sau } A_n).$$

Membrul al doilea al acestei egalități reprezintă probabilitatea evenimentului sigur și prin urmare este egal cu unu; prin urmare, pentru sistemul complect

$$P(A_1)+P(A_2)+\ldots+P(A_n)=1,$$

ceeace era de demonstrat.

Exemplul 1. La fiecare 100 de focuri trase în panoul arătat în fig. 1 (pag. 11), trăgătorul dă în medie

44 nimeriri în regiunea *1*,
30 „ „ „ *2*,
15 „ „ „ *3*,
6 „ „ „ *4*,
4 „ „ „ *5*,
1 „ „ „ *6*.

(44+30+15+6+4+1=100). Aceste 6 rezultate ale tragerii reprezintă în mod evident un sistem complect de evenimente. Probabilitățile lor sunt respectiv egale cu:

0,44; 0,30; 0,15; 0,06; 0,04; 0,01.

avem:

$$0{,}44+0{,}30+0{,}15+0{,}06+0{,}04+0{,}01=1.$$

Gloanțele care nimeresc în regiunea *6* nu nimeresc deloc sau numai parțial în panou și deci nu se pot socoti; totuși, se poate găsi probabilitatea nimeririi în această regiune, și anume, este suficient să scădem din unu suma probabilităților nimeririi în toate celelalte regiuni.

Exemplul 2. Statistica arată că într'o țesătorie oarecare la 100 de cazuri de oprire ale războiului, care necesită intervenții ulterioare ale țesătoarei, în medie:

22 au loc din cauza ruperii firelor urzelii,
31 „ „ „ „ „ „ bătăturii,
27 „ „ „ „ schimbului suveicii,
3 „ „ „ „ stricăciunilor la picker,

iar celelalte opriri ale războiului se datoresc unor cauze diverse.

Scanned with OKEN Scanner

Vedem că, în afară de cauze diverse, există patru cauze determinate ale opririi războiului, ale căror probabilități sunt respectiv egale cu:

$$0{,}22;\quad 0{,}31;\quad 0{,}27;\quad 0{,}03.$$

Suma acestor probabilități este egală cu 0,83. Împreună cu cauzele diverse, cauzele arătate ale opririi războiului formează un sistem complect de evenimente; prin urmare, probabilitatea opririi războiului din cauze diverse este egală cu:

$$1-0{,}83=0{,}17.$$

§ 6. Exemple

Pe teorema pe care am stabilit-o în legătură cu un sistem complect de evenimente, se bazează deseori și cu succes așa numitul calcul „a priori" — calculul probabilităților înainte de experiență. Să presupunem, de exemplu, că un mic teren de formă dreptunghiulară (fig. 2) împărțit în șase pătrate egale, numerotate pe figură, este supus unui foc de artilerie. Se întâmplă deseori ca pătratele care ne interesează să se afle în aceleași condiții sub focul artileriei și să nu existe niciun temei de a presupune că un anumit pătrat din cele șase pătrate va fi atins mai mult decât altul; admitem deci în mod firesc, că, în medie, fiecare dintre aceste șase pătrate va fi atins la fel de frecvent, ceeace înseamnă că probabilitățile p_1, p_2, p_3, p_4, p_5, p_6 ale lovirii acestor pătrate sunt egale între ele. Dacă se presupune că la fiecare foc tras este atins neapărat unul din pătrate, rezultă că fiecare dintre numerele p este egal cu $\frac{1}{6}$ deoarece aceste numere sunt egale între ele iar suma lor, în virtutea teoremei demonstrate, este egală cu unu. Desigur că acest rezultat, bazat pe o serie de ipoteze, cere pentru confirmarea sa o verificare experimentală; suntem însă atât de obișnuiți în asemenea cazuri ca rezultatul experiențelor să fie pozitiv, încât, bazați pe practică, putem considera adevărat ceeace am admis în mod teoretic chiar și înainte de verificarea experimentală. Se spune de obicei în asemenea cazuri că operația dată poate să aibă n rezultate diferite *egal probabile* (astfel, în cazul nostru, tragerea în terenul arătat în figura 2 poate avea ca rezultat

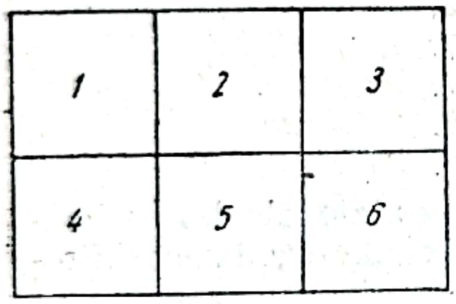

Fig. 2

Scanned with OKEN Scanner

atingerea unuia din cele șase pătrate). Probabilitatea fiecăruia în parte dintre aceste n rezultate este în acest caz egală cu $\frac{1}{n}$. Importanța unor astfel de calcule „a priori" constă în aceea că ele ne permit, în multe cazuri, să prevedem probabilitatea unui eveniment în condiții în care efectuarea unor operații de masă este cu totul imposibilă sau extrem de dificilă.

Exemplul 1. In obligațiunile Imprumutului nostru de Stat numerele seriilor sunt exprimate de obicei prin numere cu cinci cifre. Să căutăm, de exemplu, probabilitatea ca ultima cifră dintr'o serie luată la întâmplare să fie egală cu 7 (ca, de exemplu, în seria Nr. 59 607). Conform definiției pe care am dat-o probabilității, trebue să cercetăm pentru aceasta o serie lungă de tabele de tragere și să socotim câte din seriile câștigătoare au numere terminate cu cifra 7; raportul dintre acest număr și numărul total al seriilor câștigătoare reprezintă probabilitatea căutată. Suntem pe deplin îndreptățiți să presupunem că fiecare dintre cele 10 cifre 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 poate ocupa ultimul loc în numărul seriei câștigătoare, cu aceeași probabilitate ca și oricare alta dintre ele. Putem deci să presupunem, fără nicio ezitare, că probabilitatea căutată este 0,1. Puteți verifica ușor justețea acestei „prevederi" teoretice: faceți calculele necesare în limitele unei tabele de tragere oarecare și vă veți convinge că, în adevăr, fiecare din cele 10 cifre va ocupa ultimul loc în aproximativ 1/10 din toate cazurile.

Exemplul 2. Linia telefonică care unește două puncte A și B, situate la 2 km distanță unul de altul, s'a rupt într'un loc necunoscut. Care este probabilitatea ca firul să se fi rupt la o distanță de cel mult 450 m de punctul A? Descompunând, în închipuire, întreaga linie în metri, putem, datorită omogeneității reale a tuturor acestor fragmente, să admitem că probabilitatea ruperii este aceeași pentru fiecare metru. De aici, ca și în cazul precedent, rezultă imediat că probabilitatea căutată este egală cu:

$$\frac{450}{2\,000} = 0{,}225.$$

CAPITOLUL III

PROBABILITĂȚI CONDIȚIONATE ȘI REGULA DE ÎNMULȚIRE

§ 7. Noțiunea de probabilitate condiționată

Două fabrici produc becuri electrice și una din ele furnizează 70%, iar a doua 30% din totalul producției destinate consumului. Din fiecare 100 becuri produse de prima fabrică în medie 83 sunt conforme standardului [1], iar din 100 becuri produse de a doua fabrică — numai 63.

E ușor de calculat pe baza acestor date că, în medie, la 100 becuri electrice cumpărate de consumatori, revin 77 becuri conforme standardului [2], deci probabilitatea de a cumpăra un bec standard este egală cu 0,77. Să presupunem acum că am stabilit că toate becurile aflate în magazin sunt produse de prima fabrică. Probabilitatea ca becul să fie standard se schimbă în acest caz și va fi egală cu:

$$\frac{83}{100}=0{,}83.$$

Exemplul considerat arată că, dacă la condițiile generale ale operațiilor (în cazul nostru — cumpărarea unui bec electric) se adaugă o nouă condiție esențială (în exemplul nostru — stabilirea fabricii producătoare), aceasta poate schimba probabilitatea rezultatului operației individuale. Acest lucru este firesc: însăși definiția noțiunii de probabilitate cere ca ansamblul condițiilor în care are loc operația de masă să fie determinat în mod precis. Adăugând acestor condiții o condiție nouă oarecare, în general modificăm în mod esențial ansamblul acestor condiții. După aceasta, operația noastră de masă se va desfășura în condiții noi; de fapt, ea devine o altă operație, deaceea și probabilitatea unui rezultat sau altuia nu mai este aceeași ca în condițiile inițiale.

Avem astfel două probabilități ale unuia și aceluiași eveniment — cumpărarea unui bec standard — dar aceste probabilități sunt calculate în condiții diferite: dacă nu impunem o condiție suplimentară (nu ținem seama unde a fost fabricat becul) avem *o probabilitate necondiționată* egală cu 0,77 de a cumpăra un bec standard; când însă impunem o condiție suplimentară (aceea ca becul să fi fost fabricat de prima fabrică) obținem *o probabilitate*

---

[1] Vom spune că becul electric este standard (conform standardului) dacă el poate arde cel puțin 1 200 ore; în caz contrar, vom spune că becul nu este standard.

[2] In adevăr $0{,}83 \cdot 70+0{,}63 \cdot 30=77$.

Scanned with OKEN Scanner

condiționată de 0,83, care diferă în oarecare măsură de probabilitatea precedentă. Dacă însemnăm prin A evenimentul — cumpărarea unui bec standard și prin B evenimentul — fabricarea lui de prima fabrică, se însemnează în mod obișnuit prin $P(A)$ probabilitatea necondiționată a evenimentului A și prin $P_B(A)$ probabilitatea producerii aceluiași eveniment cu condiția ca să se fi produs și evenimentul B, ca becul să fi fost produs de prima fabrică. Avem astfel:

$$P(A)=0{,}77; \quad P_B(A)=0{,}83.$$

Deoarece putem vorbi despre probabilitatea unui anumit rezultat al unei operații date numai în anumite condiții bine determinate, riguros vorbind orice probabilitate este o probabilitate condiționată; nu pot exista probabilități necondiționate în sensul strict al acestui cuvânt. Totuși, în majoritatea problemelor concrete, lucrurile se prezintă astfel, că la baza tuturor operațiilor considerate în problema dată stă un anumit ansamblu determinat de condiții K, pe care le presupunem îndeplinite în toate operațiile. Dacă la calculul unei probabilități oarecare, în afară de ansamblul K nu se impune nicio altă condiție, vom numi o astfel de probabilitate *necondiționată*. Vom numi probabilitate *condiționată* probabilitatea care a fost calculată în ipoteza că în afară de ansamblul K de condiții, comune pentru toate operațiile, mai sunt îndeplinite și alte condiții, precizate în condițiile suplimentare.

Astfel, în exemplul nostru noi presupunem desigur că fabricarea becului electric se face în anumite condiții determinate, unele și aceleași pentru toate becurile puse în vânzare. Această ipoteză se impune în mod evident, deaceea în formularea problemei nici nu am găsit necesar să o menționăm. Dacă pentru un bec dat, nu impunem nicio condiție suplimentară, probabilitatea unui rezultat sau altuia al încercării acestui bec o considerăm necondiționată; iar dacă, pe lângă aceste condiții, impunem anumite restricții suplimentare, probabilitățile calculate ținând seama de aceste restricții vor fi condiționate.

Exemplul 1. În problema pe care am pus-o la începutul acestui paragraf, este evident că probabilitatea fabricării becului de către cea de a doua fabrică este egală cu 0,3. S'a stabilit că becul este de calitate standard. După această constatare, care este probabilitatea ca becul să fie fabricat de a doua fabrică?

Dintr'o mie de becuri puse în vânzare, 770 sunt de calitate standard, dintre acestea 581 sunt becuri produse de fabrica întâia

Scanned with OKEN Scanner

și 189 de cea de a doua [1]. După această observație, probabilitatea fabricării becului de către cea de a doua fabrică este deci:

$$\frac{189}{770} \approx 0{,}245.$$

Aceasta este probabilitatea condiționată ca becul să fi fost produs de fabrica a doua, în ipoteza că becul dat este standard. Cu notațiile precedente, putem scrie:

$$P(B) = 0{,}3; \quad P_A(\bar{B}) \approx 0{,}245$$

(evenimentul $\bar{B}$ reprezintă nerealizarea evenimentului B).

Exemplul 2. Observațiile făcute în curs de mai mulți ani într'un raion oarecare au arătat că, dintr'o sută de mii de copii care au atins vârsta de 10 ani, până la vârsta de 40 ani ajung în medie 82 277, iar până la 70 ani — 37 977. Să se găsească probabilitatea ca un om care a atins vârsta de 40 ani să trăiască până la vârsta de 70 ani.

Deoarece din 82 277 persoane de 40 ani, în medie 37 977 ajung până la 70 ani, probabilitatea ca o persoană de 40 ani să ajungă la vârsta de 70 ani este egală cu:

$$\frac{37\,977}{82\,277} \approx 0{,}46.$$

Dacă notăm cu A evenimentul care constă în faptul că un copil de 10 ani trăiește până la 70 ani și cu B — că el trăiește până la 40 de ani, este evident că avem:

$$P(A) = 0{,}37977 \approx 0{,}38; \quad P_B(A) \approx 0{,}46.$$

§ 8. Deducerea regulii de înmulțire a probabilităților

Să revenim la primul exemplu din paragraful precedent. Pentru fiecare 1 000 becuri vândute, 300 în medie sunt produse de a doua fabrică, iar din aceste 300 becuri 189 sunt de calitate standard. Vedem deci că probabilitatea ca becul să fi fost fabricat de a doua fabrică (evenimentul $\bar{B}$) este egală cu:

$$P(\bar{B}) = \frac{300}{1\,000} = 0{,}3,$$

[1] Aceasta se poate calcula ușor în felul următor: pentru fiecare 1 000 becuri, 700 (în medie) sunt fabricate de prima fabrică, și, la fiecare 100 becuri din prima fabrică, 83 (în medie) au calitate standard. Deci pentru 700 becuri din prima fabrică $7 \cdot 83 = 581$ becuri în medie sunt de calitate standard. Celelalte 189 becuri de calitate standard sunt produse de fabrica a doua.

Scanned with OKEN Scanner

iar probabilitatea ca el să fie de calitate standard, dacă a fost produs de a doua fabrică, este egală cu:

$$P_{\bar{B}}(A)=\frac{189}{300}=0{,}63.$$

Deoarece pentru fiecare 1 000 becuri, 189 sunt produse de fabrica a doua și sunt totodată de calitate standard, probabilitatea producerii simultane a evenimentelor A și $\bar{B}$ este egală cu:

$$P(A \text{ și } \bar{B})=\frac{189}{1\,000}=\frac{300}{1\,000}\cdot\frac{189}{300}=P(\bar{B})\cdot P_{\bar{B}}(A).$$

Această „regulă de înmulțire" se poate ușor extinde și la cazuri generale. Să presupunem că, în fiecare serie de n operații, evenimentul B se produce în medie de m ori, iar în fiecare serie de m asemenea operații, în care s'a produs evenimentul B, evenimentul A se produce de l ori. In acest caz, în fiecare serie de n operații producerea simultană a evenimentelor B și A se va observa în medie de l ori. Astfel:

$$P(B)=\frac{m}{n},\quad P_B(A)=\frac{l}{m},$$

$$P(A \text{ și } B)=\frac{l}{n}=\frac{m}{n}\cdot\frac{l}{m}=P(B)\cdot P_B(A). \qquad (1)$$

Regula de înmulțire. *Probabilitatea producerii simultane a două evenimente este egală cu produsul dintre probabilitatea primului eveniment și probabilitatea condiționată a celui de al doilea, calculată în ipoteza că primul eveniment a avut loc.*

Putem, desigur, considera ca primul, oricare din aceste două evenimente, astfel încât formula (1) se poate scrie tot atât de bine sub forma:

$$P(A \text{ și } B)=P(A)\cdot P_A(B), \qquad (1')$$

de unde obținem o relație importantă:

$$P(A)\cdot P_A(B)=P(B)\cdot P_B(A) \qquad (2)$$

Am avut în exemplul nostru:

$$P(A \text{ și } \bar{B})=\frac{189}{1\,000},\quad P(A)=\frac{77}{100},\quad P_A(\bar{B})=\frac{189}{770};$$

ceeace confirmă formula (1').

Exemplu. La o întreprindere oarecare, 96% din produse sunt calificate satisfăcătoare (evenimentul A); 75 în medie dintr'o sută de produse satisfăcătoare sunt de calitatea I-a (evenimentul B). Să se afle probabilitatea ca un produs fabricat de această întreprindere să fie de calitatea I-a.

Scanned with OKEN Scanner

Se caută $P(A$ și $B)$, deoarece pentruca un produs să fie de calitatea I-a trebue să fie satisfăcător (evenimentul A) și de calitatea I-a (evenimentul B).

Conform condițiilor problemei

$$P(A)=0{,}96, \quad P_A(B)=0{,}75.$$

Deacecea, în baza formulei (1')

$$P(A \text{ și } B)=0{,}96\cdot 0{,}75=0{,}72.$$

§ 9. Evenimente independente

La încercarea rezistenței a două mosoare de fire de bumbac, fabricate la mașini diferite, s'a constatat că, pentru primul mosor, o mostră din el de o anumită lungime rezistă la un efort standard determinat, cu probabilitatea de 0,84, iar o mostră dintr'al doilea mosor — cu probabilitatea de 0,78 [1]. Să se afle probabilitatea pentruca două mostre de fire luate din două mosoare diferite să fie amândouă capabile să reziste la efortul standard.

Să însemnăm prin A evenimentul care constă în faptul că o mostră luată din primul mosor rezistă la efortul standard și să însemnăm prin B evenimentul analog pentru o mostră din cel de al doilea mosor. Deoarece căutăm pe $P(A$ și $B)$, aplicăm regula înmulțirii:

$$P(A \text{ și } B)=P(A)\cdot P_A(B).$$

$P(A)$ în acest caz este egală, evident, cu 0,84; dar care este valoarea lui $P_A(B)$? Conform definiției generale a probabilităților condiționate, $P_A(B)$ reprezintă probabilitatea ca mostra de fire dintr'al doilea mosor să reziste la efortul standard, dacă același efort a fost suportat de mostra din primul mosor. Probabilitatea evenimentului B nu depinde însă de faptul că evenimentul A a avut sau nu loc, cu atât mai mult cu cât putem face aceste încercări concomitent, iar mostrele de fire sunt luate din mosoare cu totul diferite, fabricate la mașini diferite. Practic, aceasta înseamnă că procentul încercărilor în care fire din cel de al doilea mosor vor rezista la efortul standard nu depinde de rezistența opusă de mostra din primul mosor, adică:

$$P_A(B)=P(B)=0{,}78;$$

de unde rezultă că:

$$P(A \text{ și } B)=P(A)\cdot P(B)=0{,}84 \cdot 0{,}78=0{,}6552.$$

[1] Dacă efortul standard este, de pildă, de 400 gr, aceasta înseamnă că din 100 mostre luate din primul mosor, 84 în medie rezistă la efort, iar 16 nu rezistă și se rup.

Scanned with OKEN Scanner

Particularitatea prin care acest exemplu diferă de toate cele precedente constă, după cum se vede, în faptul că aici, probabilitatea rezultatului B nu se schimbă, atunci când la condițiile comune mai adăugăm cerința ca evenimentul A să aibă loc. Cu alte cuvinte, probabilitatea condiționată $P_A(B)$ este egală cu probabilitatea necondiționată $P(B)$. In acest caz, vom spune pe scurt că *evenimentul B nu depinde de evenimentul A.*

Ne putem convinge ușor că, dacă B nu depinde de A, nici A nu depinde de B; în adevăr, dacă $P_A(B)=P(B)$, în virtutea formulei (2) și $P_B(A)=P(A)$, și aceasta exprimă tocmai faptul că evenimentul A nu depinde de evenimentul B. Astfel, independența a două evenimente este o proprietate *reciprocă*. Vedem că, pentru evenimente reciproc independente, regula de înmulțire capătă o formă extrem de simplă:

$$P(A \text{ și } B)=P(A)\cdot P(B). \qquad (3)$$

După cum la fiecare caz de aplicare a regulii de adunare, este necesar să se stabilească în prealabil incompatibilitatea reciprocă a evenimentelor date, în fiecare caz de aplicare a regulii (3) trebue să ne convingem că evenimentele A și B sunt reciproc independente. Neglijarea acestei verificări duce la multe greșeli. Dacă evenimentele A și B sunt reciproc dependente, formula (3) nu este justă și trebue înlocuită cu formula mai generală (1) sau (1').

Regula (3) se extinde ușor la cazul în care se caută probabilitatea producerii nu a două ci a trei și mai multe evenimente reciproc independente. Să presupunem, de exemplu, că avem trei evenimente reciproc independente A, B și C (aceasta înseamnă că probabilitatea fiecăruia dintre ele nu depinde de faptul că celelalte două evenimente au avut loc sau nu). Deoarece evenimentele A, B și C sunt reciproc independente, conform regulii (3):

$$P(A \text{ și } B \text{ și } C)=P(A \text{ și } B)\cdot P(C);$$

introducând aici în locul lui $P(A$ și $B)$ expresia acestei probabilități, dată de formula (3), obținem:

$$P(A \text{ și } B \text{ și } C)= P(A)\cdot P(B)\cdot P(C), \qquad (4)$$

și este clar că aceeași regulă este valabilă în cazul în care grupul considerat conține un număr oarecare de evenimente, cu condiția însă ca aceste evenimente să fie reciproc independente (adică probabilitatea fiecăruia dintre ele să nu depindă de producerea sau neproducerea celorlalte evenimente).

Probabilitatea producerii simultane a unui număr oarecare de evenimente reciproc independente este egală cu produsul probabilităților acestor evenimente.

Scanned with OKEN Scanner

Exemplul 1. Un muncitor deservește trei stunguri. Probabilitatea ca timp de o oră strungul să nu ceară îngrijirea muncitorului este pentru primul strung egală cu 0,9, pentru al doilea 0,8 și pentru al treilea 0,85. Să se afle probabilitatea ca timp de o oră niciunul din strunguri să nu aibă nevoie de îngrijirea muncitorului.

Presupunând că strungurile lucrează independent unul de altul și aplicând formula (4) se obține probabilitatea căutată egală cu:

$$0{,}9 \cdot 0{,}8 \cdot 0{,}85 = 0{,}612.$$

Exemplul 2. In condițiile exemplului 1, să se afle probabilitatea ca cel puțin unul din cele trei strunguri să nu aibă nevoie, timp de o oră, de îngrijirea muncitorului.

In această problemă, intervine o probabilitate de forma $P(A$, sau B, sau $C)$, deaceea raționamentul nostru ne duce în mod firesc către regula de adunare. Ne putem convinge însă imediat că, în acest caz, regula de adunare nu este aplicabilă, întrucât oricare două din cele trei evenimente considerate sunt compatibile între ele (nimic nu împiedică două strunguri să funcționeze neîntrerupt în timpul aceleiași ore); dar, chiar independent de acest raționament, este evident că suma celor trei probabilități date este mult mai mare decât unu, deci această sumă nu poate fi o probabilitate.

Pentru rezolvarea problemei, să observăm că probabilitatea ca primul din strunguri să necesite atenția și îngrijirea muncitorului este egală cu 0,1, pentru al doilea strung 0,2 și pentru al treilea strung ea este 0,15. Intrucât aceste trei evenimente sunt reciproc independente, probabilitatea producerii tuturor acestor trei evenimente, conform regulii (4), este egală cu:

$$0{,}1 \cdot 0{,}2 \cdot 0{,}15 = 0{,}003.$$

Dar evenimentele „toate trei strunguri vor avea nevoie de îngrijire" și „cel puțin unul dintre ele va funcționa neîntrerupt", sunt, evident, două evenimente contrarii. Suma probabilităților lor este deci egală cu 1 și, prin urmare, probabilitatea căutată este egală cu $1-0{,}003=0{,}997$. Când probabilitatea unui eveniment este atât de apropiată de unu, acest eveniment poate fi socotit în mod practic ca sigur. Aceasta înseamnă că aproape întotdeauna, în timp de o oră, cel puțin unul dintre cele trei strunguri va funcționa neîntrerupt.

Exemplul 3. In anumite condiții determinate, probabilitatea de a dobori un avion al dușmanului cu un foc tras dintr'o pușcă este egală cu 0,004. Să se găsească probabilitatea doboririi avionului dușman printr'o tragere simultană din 250 puști.

In cazul unei singure trageri există o probabilitate de $1-0{,}004=0{,}996$ ca avionul să nu fie doborît. Probabilitatea ca

Scanned with OKEN Scanner

el să nu fie doborît de toate cele 250 de focuri trase este, conform regulii de înmulțire a probabilităților evenimentelor independente, egală cu produsul a 250 factori, egali cu 0,996, adică această probabilitate este egală cu $(0{,}996)^{250}$. Iar probabilitatea ca una cel puțin din aceste 250 de împușcături să doboare avionul este deci egală cu:

$$1-(0{,}996)^{250}.$$

Un calcul amănunțit, pe care nu-l vom reproduce aici, arată că acest număr este egal aproximativ cu $\frac{5}{8}$. Astfel, cu toate că probabilitatea de a dobori un avion dușman cu un singur foc de pușcă este foarte mică (0,004), la o tragere simultană dintr'un număr de puști, probabilitatea unui rezultat util devine mult mai însemnată.

Raționamentul făcut în ultimele două exemple poate fi ușor generalizat și duce la o importantă regulă generală. In ambele exemple a fost vorba de probabilitatea $P(A_1$, sau A_2, sau ..., $A_n)$ a producerii unuia sau câtorva dintre evenimentele reciproc independente $A_1, A_2, \ldots, A_n$. Dacă însemnăm prin $\bar{A}_k$ evenimentul care constă în nerealizarea lui A_k, evenimentele A_k și $\bar{A}_k$ sunt contrarii, astfel încât:

$$P(A_k)+P(\bar{A}_k)=1.$$

Pe de altă parte, evenimentele $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n$, sunt, evident, reciproc independente, deci:

$$P(\bar{A}_1 \text{ și } \bar{A}_2 \text{ și } \ldots \text{ și } \bar{A}_n)=P(\bar{A}_1)\cdot P(\bar{A}_2)\cdot \ldots \cdot P(\bar{A}_n)=$$
$$=[1-P(A_1)]\cdot[1-P(A_2)]\cdot \ldots \cdot[1-P(A_n)].$$

In sfârșit, evenimentele (A_1 sau A_2 sau ... sau A_n) și ($\bar{A}_1$ și $\bar{A}_2$ și ... și $\bar{A}_n$) sunt, după cum se vede, contrarii între ele (există două posibilități: sau se produce cel puțin unul din evenimentele A_k sau se produc toate evenimentele $\bar{A}_k$). De aceea:

$$P(A_1 \text{ sau } A_2 \text{ sau } \ldots \text{ sau } A_n)=1-P(\bar{A}_1 \text{ și } \bar{A}_2 \text{ și } \ldots \text{ și } \bar{A}_n)=$$
$$=1-[1-P(A_1)]\cdot[1-P(A_2)]\cdot \ldots \cdot[1-P(A_n)]. \qquad (5)$$

Aceasta este o formulă importantă, cu ajutorul căreia putem calcula probabilitatea producerii *cel puțin a unuia* dintre evenimentele $A_1, A_2, \ldots, A_n$, cunoscând probabilitățile acestor evenimente, formula fiind valabilă atunci și numai atunci când aceste evenimente sunt reciproc independente. Dacă, în particular, toate evenimentele A_k au una și aceeași probabilitate p (așa cum s'a întâmplat în exemplul 3),

$$P(A_1 \text{ sau } A_2 \text{ sau } \ldots \text{ sau } A_n)=1-(1-p)^n. \qquad (6)$$

Scanned with OKEN Scanner

Exemplul 4. La o tragere într'o țintă aeriană, focul tras este considerat în anumite condiții ca eficace dacă punctul unde a explodat proectilul este situat la cel mult 10 m în fiecare din cele trei direcții (lungime, lățime, înălțime) de centrul țintei; cu alte cuvinte, pentru eficacitatea tragerii proectilul trebue să facă explozie în interiorul unui cub, având muchia egală cu 20 m și al cărui centru coincide cu centrul țintei. Probabilitatea abaterilor care depășesc 10 m este:

în lungime $p_1=0{,}08$,
în lățime $p_2=0{,}12$,
în înălțime $p_3=0{,}1$;

să se afle probabilitatea ca focul tras să fie ineficace.

Pentru ca focul să fie ineficace, trebue ca în cel puțin una din cele trei direcții abaterea locului de explozie al proectilului dela centrul țintei să depășească 10 m; deoarece aceste trei evenimente pot fi considerate deobicei ca reciproc independente (fiind provocate în fond de cauze diferite) pentru rezolvarea acestei probleme putem folosi formula (5); astfel încât:

$$P=1-(1-p_1)\cdot(1-p_2)\cdot(1-p_3)\approx 0{,}27.$$

Putem deci presupune că, din 100 focuri, aproximativ 73 vor fi eficace.

CAPITOLUL IV

CONSECINȚE ALE REGULILOR DE ADUNARE ȘI ÎNMULȚIRE

§ 10. Stabilirea câtorva inegalități

Să revenim la exemplul cu becurile electrice din capitolul precedent (pag. 18). Să introducem următoarele notații pentru evenimente:

A — becul electric de calitate standard,
$\overline{A}$ — becul electric de calitate neconformă cu standardul,
B — becul electric produs de fabrica întâia,
$\overline{B}$ — becul electric produs de fabrica a doua.

Evident că evenimentele A și $\overline{A}$ reprezintă două evenimente contrarii; tot așa și evenimentele B și $\overline{B}$ sunt două evenimente contrarii.

Scanned with OKEN Scanner

Dacă becul este standard (A), el poate fi produs sau de fabrica întâia (A și B) sau de cea de a doua ($\bar{A}$ și $\bar{B}$); deoarece ultimele două evenimente sunt în mod evident incompatibile între ele, avem conform regulii de adunare:

$$P(A)=P(A \text{ și } B)+P(A \text{ și } \bar{B}). \qquad (1)$$

In mod analog obținem:

$$P(B)=P(A \text{ și } B)+P(\bar{A} \text{ și } B). \qquad (2)$$

In sfârșit să examinăm evenimentul (A sau B); pentru producerea lui există, evident, următoarele trei posibilități:

1. A și B,
2. A și $\bar{B}$,
3. $\bar{A}$ și B.

Oricare două din aceste trei posibilități sunt incompatibile între ele; deaceea, conform teoremei de adunare

$$P(A \text{ sau } B)=P(A \text{ și } B)+P(A \text{ și } \bar{B})+P(\bar{A} \text{ și } B). \qquad (3)$$

Adunând egalitățile (1) și (2) membru cu membru și ținând seamă de egalitatea (3), obținem cu ușurință:

$$P(A)+P(B)=P(A \text{ și } B)+P(A \text{ sau } B),$$

de unde

$$P(A \text{ sau } B)=P(A)+P(B)-P(A \text{ și } B). \qquad (4)$$

Am ajuns astfel la un rezultat foarte important; și, cu toate că am raționat pe un exemplu particular, el a fost suficient de general pentru ca să putem considera concluziile trase valabile pentru orice pereche de evenimente A și B. Până acum am obținut expresiile pentru probabilitatea $P(A \text{ sau } B)$ numai în ipoteze particulare, privind relațiile dintre evenimentele A și B (la început le-am considerat incompatibile, iar apoi independente între ele). Formula (4) la care am ajuns este valabilă, fără nicio ipoteză suplimentară, pentru orice pereche de evenimente A și B. Nu trebue însă să uităm o deosebire esențială care există între formula (4) și formulele noastre anterioare. In formulele anterioare, probabilitatea $P(A \text{ sau } B)$ se exprima întotdeauna numai cu ajutorul probabilităților $P(A)$ și $P(B)$, astfel încât, cunoscând probabilitățile evenimentelor A și B, puteam determina întotdeauna în mod univoc probabilitatea evenimentului (A sau B). Altfel se prezintă lucrurile în formula (4): pentru a calcula cu această formulă mărimea $P(A \text{ sau } B)$, trebue ca, pe lângă $P(A)$ și $P(B)$, să cunoaștem și $P(A \text{ și } B)$, adică probabilitatea producerii simultane a evenimentelor A și B; dar găsirea acestei probabilități într'un caz general, adică pentru relații arbitrare între evenimentele A

Scanned with OKEN Scanner

și B, este deobicei tot atât de anevoioasă ca și găsirea lui $P(A$ sau $B)$; deaceea formula (4) este rar folosită în calculele practice directe, dar ea are o mare importanță teoretică:

Să ne convingem în primul rând că putem obține ușor din formula (4), ca niște cazuri particulare, formulele anterioare. Dacă evenimentele A și B sunt incompatibile între ele, evenimentul (A și B) este imposibil, $P(A$ și $B)=0$, și formula (4) devine

$$P(A \text{ sau } B)=P(A)+P(B),$$

adică regula de adunare. Dacă evenimentele A și B sunt reciproc independente, avem conform formulei (3) din pag. 23

$$P(A \text{ și } B)=P(A)\cdot P(B),$$

și formula (4) ne dă:

$$P(A \text{ sau } B)=P(A)+P(B)-P(A)\cdot P(B)=$$
$$=1-[1-P(A)]\cdot[1-P(B)],$$

adică formula (5) din pag. 25 (pentru cazul $n=2$).

Vom deduce acum din formula (4) un rezultat important. Deoarece în orice caz $P(A$ și $B)\geqq 0$ din formula (4) rezultă că în orice caz

$$P(A \text{ sau } B)\leq P(A)+P(B). \qquad (5)$$

Această inegalitate poate fi lesne extinsă la un număr oarecare de evenimente. Astfel, în cazul a trei evenimente, avem conform formulei (5),

$$P(A \text{ sau } B \text{ sau } C)\leq P(A \text{ sau } B)+P(C)\leq P(A)+P(B)+P(C),$$

și, în același mod, putem evident să trecem dela 3 evenimente la 4, etc. Obținem deci un rezultat general.

Probabilitatea producerii a cel puțin unul din mai multe evenimente nu depășește niciodată suma probabilităților acestor evenimente.

Semnul egalității are loc numai atunci când evenimentele sunt incompatibile două câte două.

§ 11. Formula probabilității totale

Să revenim încăodată la exemplul cu becurile din pag. 18 și să folosim pentru diversele rezultate ale experiențelor notațiile introduse la pagina 20. Probabilitatea ca becul să aibă calitate standard, cu condiția ca să fie fabricat de a doua fabrică este, după cum am mai văzut la pag. 21, egală cu:

$$P_{\bar{B}}(A)=\frac{189}{300}=0{,}63;$$

Scanned with OKEN Scanner

probabilitatea aceluiași eveniment cu condiția ca becul să fi fost produs de prima fabrică este:

$$P_B(A)=\frac{581}{700}=0{,}83.$$

Să presupunem că aceste două numere ne sunt cunoscute și că cunoaștem și probabilitatea ca becul să fie fabricat de prima întreprindere

$$P(B)=0{,}7$$

sau de către cea de a doua întreprindere

$$P(\bar{B})=0{,}3.$$

Trebue să aflăm probabilitatea necondiționată $P(A)$, adică probabilitatea calității standard a unui bec oarecare, calculată fără niciun fel de condiție, privind locul ei de fabricație.

Pentru a rezolva această problemă, vom raționa în felul următor: însemnăm prin E evenimentul dublu care constă într'aceea că: 1. becul este fabricat de prima întreprindere și 2. becul este standard; iar prin F — același eveniment pentru a doua fabrică. Deoarece orice bec standard este produs de prima sau a doua fabrică, evenimentul A este echivalent cu evenimentul «E sau F», iar deoarece evenimentele E și F sunt incompatibile între ele, conform regulii de adunare obținem:

$$P(A)=P(E)+P(F). \qquad (6)$$

Pe de altă parte, pentru ca să aibă loc evenimentul E, trebue: 1. ca becul să fie produs de prima fabrică (B) și 2. becul să fie standard (A): evenimentul E este deci echivalent cu evenimentul «B și A», de unde, conform regulii de înmulțire, rezultă:

$$P(E)=P(B)\cdot P_B(A).$$

In mod absolut analog obținem:

$$P(F)=P(\bar{B})\cdot P_B(A),$$

și, introducând aceste expresii în egalitatea (6), avem:

$$P(A)=P(B)\cdot P_B(A)+P(\bar{B})\cdot P_{\bar{B}}(A).$$

Această formulă dă soluția problemei pe care ne-am pus-o. Introducând datele numerice, găsim că $P(A)=0{,}77$.

Exemplu. Pentru însămânțări s'au pregătit semințe de grâu de calitatea I-a, conținând însă și cantități neinsemnate de alte calități — II, III și IV. Să luăm un bob oarecare din aceste grăunțe. Vom însemna prin A_1 evenimentul care constă în faptul că bobul este de calitatea I-a, prin A_2 — că el este de calitatea a II-a, prin A_3 — că este de calitatea a III-a, și în sfârșit prin A_4

Scanned with OKEN Scanner

— că este de calitatea a IV-a. Probabilităţile ca un bob oarecare, luat la întâmplare, să fie din una din aceste calităţi ne sunt cunoscute, ele sunt respectiv egale cu:

$$P(A_1)=0{,}96;\quad P(A_2)=0{,}01;\quad P(A_3)=0{,}02;\quad P(A_4)=0{,}01.$$

(Suma acestor patru numere este egală cu unu, aşa cum trebue să fie în cazul unui sistem complect de evenimente). Probabilitatea ca dintr'un bob să crească un spic care să conţină cel puţin 50 boabe este egală cu:

1. 0,50 pentru bobul calitatea I,
2. 0,15 „ „ „ II,
3. 0,20 „ „ „ III,
4. 0,05 „ „ „ IV.

Se cere să se calculeze probabilitatea necondiţionată ca spicul să aibă cel puţin 50 boabe.

Să însemnăm prin K evenimentul constând din faptul că spicul va conţine cel puţin 50 boabe; atunci, conform condiţiilor problemei

$$P_{A_1}(K)=0{,}50;\quad P_{A_2}(K)=0{,}15;\quad P_{A_3}(K)=0{,}20;\quad P_{A_4}(K)=0{,}05.$$

Noi trebue să determinăm pe $P(K)$. Să notăm cu E_1 evenimentul constând în faptul că bobul este de calitatea A_1, şi că spicul crescut din el conţine cel puţin 50 de boabe, astfel încât E_1 este echivalent cu evenimentul (A_1 şi K). In mod analog, vom nota cu

E_2 — evenimentul (A_2 şi K),
E_3 — „ (A_3 şi K),
E_4 — „ (A_4 şi K).

Este evident că, pentru producerea evenimentului K, este necesar să se realizeze unul din evenimentele E_1, E_2, E_3, E_4; deoarece aceste evenimente sunt incompatibile două câte două, obţinem, conform regulii de adunare:

$$P(K)=P(E_1)+P(E_2)+P(E_3)+P(E_4). \qquad (7)$$

Pe de altă parte, din regula înmulţirii, rezultă:

$$P(E_1)=P(A_1 \text{ şi } K)=P(A_1)\cdot P_{A_1}(K),$$
$$P(E_2)=P(A_2 \text{ şi } K)=P(A_2)\cdot P_{A_2}(K).$$
$$P(E_3)=P(A_3 \text{ şi } K)=P(A_3)\cdot P_{A_3}(K),$$
$$P(E_4)=P(A_4 \text{ şi } K)=P(A_4)\cdot P_{A_4}(K),$$

Introducând aceste expresii în egalitatea (7), obţinem:

$$P(K)=P(A_1)\cdot P_{A_1}(K)+P(A_2)\cdot P_{A_2}(K)+P(A_3)\cdot P_{A_3}(K)+$$
$$+P(A_4)\cdot P_{A_4}(K);$$

şi astfel problema noastră este rezolvată.

Scanned with OKEN Scanner

Introducând datele numerice din problemă, obținem:

$$P(K)=0{,}486.$$

Cele două exemple pe care le-am examinat aici mai amănunțit ne conduc la o importantă regulă generală; acum o putem formula și demonstra fără nicio dificultate. Să presupunem că o operație dată admite rezultatele $A_1, A_2, \ldots, A_n$, care formează un sistem complect de evenimente (să reținem: aceasta înseamnă că oricare două din aceste evenimente sunt incompatibile unul cu altul și că, totodată, unul din ele trebue să se producă în mod obligatoriu). Atunci, pentru un rezultat posibil oarecare K al acestei operații, are loc relația:

$$P(K)=P(A_1)\cdot P_{A_1}(K)+P(A_2)\cdot P_{A_2}(K)+\ldots+P(A_n)\cdot P_{A_n}(K). \quad (8)$$

Această regulă (8) este numită deobicei „*formula probabilității totale*". Demonstrația acestei formule se face la fel ca și în cele două exemple examinate; în primul rând, realizarea evenimentului K cere realizarea unuia din evenimentele «A_i și K», astfel încât, conform regulii adunării,

$$P(K)=\sum_{i=1}^{n}P(A_i \text{ şi } K); \quad (9)$$

și, în al doilea rând, conform regulii de înmulțire,

$$P(A_i \text{ şi } K)=P(A_i)\cdot P_{A_i}(K);$$

introducând aceste expresii în egalitatea (9), obținem formula (8).

§ 12. Formula lui Bayes

Formulele din paragraful precedent ne dau posibilitatea să deducem un rezultat important și care are numeroase aplicații. Vom începe cu deducerea formală, amânând lămurirea definitivă a sensului real al formulei până la examinarea exemplelor.

Să considerăm din nou un sistem de evenimente $A_1, A_2, \ldots, A_n$, rezultate ale unei operații oarecare, sistem pe care îl presupunem complect. Dacă însemnăm prin K un rezultat arbitrar al acestei operații, atunci, conform regulii de înmulțire:

$$P(A_i \text{ şi } K)=P(A_i)\cdot P_{A_i}(K)=P(K)\cdot P_K(A_i) \qquad (1\leq i\leq n),$$

de unde

$$P_K(A_i)=\frac{P(A_i)\cdot P_{A_i}(K)}{P(K)} \qquad (1\leq i\leq n),$$

Scanned with OKEN Scanner

sau, exprimând numitorul fracției obținute cu ajutorul formulei probabilității totale (8) din paragraful precedent,

$$P_K(A_i) = \frac{P(A_i) \cdot P_{A_i}(K)}{\sum_{r=1}^{n} P(A_r) \cdot P_{A_r}(K)} \qquad (1 \leq i \leq n). \qquad (10)$$

Aceasta este *formula lui Bayes*, care are multe aplicații în practica calculului probabilităților. Ea este folosită cel mai des în cazuri care sunt ilustrate de următorul exemplu.

Să presupunem că se face o tragere într'o țintă situată pe o porțiune rectilinie MN (fig. 3), pe care o împărțim mintal în cinci segmente mici a, b', b'', c', c''. Presupunem că nu cunoaștem poziția exactă a țintei; știm numai care sunt probabilitățile ca ținta să se afle pe unul sau altul din cele 5 segmente; aceste probabilități sunt egale cu:

$$P(a) = 0{,}48; \quad P(b') = P(b'') = 0{,}21; \quad P(c') = P(c'') = 0{,}05,$$

unde prin a, b', b'', c', c'' însemnăm următoarele evenimente: ținta se află în segmentul a, b', b'', c', c'' (suma acestor numere este egală cu unitatea). Cea mai mare probabilitate corespunde segmentului a spre care, firește, dirijăm focul pe care urmează să-l tragem. Totuși, greșelile fiind inevitabile în tragere, ținta poate fi atinsă și atunci când ea nu se află în a, ci într'unul oarecare din celelalte segmente. Să presupunem că probabilitatea atingerii țintei (evenimentul K) este egală cu:

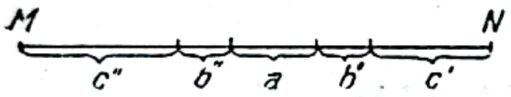

Fig. 3

$P_a(K) = 0{,}56$, dacă ținta este situată în segmentul a,
$P_{b'}(K) = 0{,}18$, „ „ „ „ „ „ b',
$P_{b''}(K) = 0{,}16$, „ „ „ „ „ „ b'',
$P_{c'}(K) = 0{,}06$, „ „ „ „ „ „ c',
$P_{c''}(K) = 0{,}02$, „ „ „ „ „ „ c''.

Să presupunem că focul s'a tras și ținta a fost atinsă (s'a produs evenimentul K). În urma acestui fapt, probabilitățile pe care le-am avut înainte pentru diverse poziții ale țintei [adică numerele $P(a)$, $P(b')$ ...], se vor schimba; latura calitativă a acestei schimbări este clară fără niciun calcul; am tras în segmentul a și am nimerit ținta — este deci clar că probabilitatea $P(a)$ trebue în acest caz să se mărească; dar noi vrem să calculăm cantitativ schimbarea produsă de focul pe care l-am tras, adică vrem să găsim expresia exactă a probabilităților $P_k(a)$, $P_k(b')$, ..., în diferitele poziții posibile ale țintei, cu condiția ca prin focul

Scanned with OKEN Scanner

tras ţinta să fi fost atinsă. Formula lui Bayes (10) ne dă imediat soluţia acestei probleme.

Astfel:

$$P_K(a)=\frac{P(a)\cdot P_a(K)}{P(a)\cdot P_a(K)+P(b')\cdot P_{b'}(K)+P(b'')\cdot P_{b''}(K)+P(c')\cdot P_{c'}(K)+P(c'')\cdot P_{c''}(K)}\approx 0{,}8;$$

constatăm deci că $P_K(a)$ este într'adevăr mai mare decât $P(a)$.

In mod analog, putem calcula cu uşurinţă şi probabilităţile $P_K(b')$, . . ., corespunzătoare altor poziţii ale ţintei. Pentru calcul este util să observăm că expresiile acestor probabilităţi date de formula lui Bayes diferă între ele numai prin numărătorii lor; numitorul este însă acelaşi pentru toate; el este egal cu $P(K)\approx 0{,}34$.

O schemă generală a unor situaţii de acest fel poate fi dată în felul următor. In condiţiile operaţiei intră un element oarecare, necunoscut, în legătură cu care se pot face n „ipoteze" diferite: $A_1, A_2, \ldots, A_n$, care formează un sistem complect de evenimente; din anumite cauze, cunoaştem probabilităţile $P(A_i)$ ale acestor ipoteze încă înainte de experienţă; cunoaştem deasemenea că, în ipoteza A_i, probabilitatea unui eveniment oarecare K (de exemplu nimerirea la ţintă) este $P_{A_i}(K)\,(1\leq i\leq n)$ [$P_{A_i}(K)$ este probabilitatea evenimentului K calculată în ipoteza că evenimentul A_i s'a produs]. Dacă, în urma experienţei, evenimentul K s'a produs, acest fapt ne obligă să evaluăm probabilităţile corespunzătoare ipotezelor A_i iar problema constă în găsirea noilor probabilităţi $P_K(A_i)$ ale acestor ipoteze şi o rezolvăm cu ajutorul formulei lui Bayes.

In practica artileriei se execută aşa numita tragere de încercare, care are ca scop să precizeze cunoştinţele noastre referitoare la condiţiile tragerii. In acest caz, elementele necunoscute care trebue precizate pot fi atât poziţia ţintei, cât şi oricare alt element din condiţiile tragerii, care influenţează eficacitatea ei (în special, o anumită particularitate a armei folosite). Se întâmplă însă, deseori, că se trag nu numai unul ci câteva focuri de încercare şi se pune atunci problema calculării noilor probabilităţi ale ipotezelor, pe baza rezultatelor obţinute în tragere. In toate aceste cazuri, problema se rezolvă uşor cu ajutorul formulei lui Bayes.

Pentru simplificare, să notăm în schema generală pe care am examinat-o

$$P(A_i)=P_i,\quad P_{A_i}(K)=p_i\qquad (1\leq i\leq n),$$

astfel încât formula lui Bayes capătă forma simplă:

$$P_K(A_i)=\frac{P_i p_i}{\sum_{r=1}^{n} P_r p_r}.$$

Scanned with OKEN Scanner

Să presupunem că s'au tras s focuri de probă și rezultatul K s'a produs de m ori și nu s'a produs de $s-m$ ori. Să notăm cu K^* rezultatul obținut dintr'o serie de s focuri trase. Putem presupune că rezultatele unor focuri diferite sunt evenimente reciproc independente. Dacă ipoteza A_i este îndeplinită, probabilitatea rezultatului K este egală cu p_i, deci probabilitatea evenimentului complementar (adică neproducerea lui K) este egală cu $1-p_i$. Probabilitatea rezultatului compus K^* este deci, conform regulii de înmulțire pentru evenimente independente, egală cu produsul a s factori, dintre care m sunt egali cu p_i, iar $s-m$ sunt egali cu $1-p_i$; cu alte cuvinte:

$$P_{A_i}(K^*)=C_s^m p_i^m (1-p_i)^{s-m} \qquad (1\leq i\leq n),$$

și conform formulei lui Bayes:

$$P_{K^*}(A_i)=\frac{P_i p_i^m (1-p_i)^{s-m}}{\sum_{r=1}^{n} P_r p_r^m (1-p_r)^{s-m}} \qquad (1\leq i\leq n), \qquad (11)$$

ceeace rezolvă problema.

Se înțelege că asemenea probleme apar nu numai în practica artileristului ci și în alte domenii de activitate.

Exemplul 1. Să se găsească, în problema pe care am studiat-o la începutul acestui paragraf, probabilitatea ca ținta să fie situată pe segmentul a, dacă două focuri trase consecutiv în acest segment au nimerit amândouă în țintă.

Insemnând prin K^* dubla nimerire la țintă, avem conform formulei (11):

$$P_{K^*}(a)=\frac{P(a)\cdot[P_a(K)]^2}{P(a)\cdot[P_a(K)]^2+P(b')\cdot[P_{b'}(K)]^2+\ldots};$$

lăsăm pe cititor să efectueze câteva calcule și să se convingă că, în urma nimeririi de două ori în țintă, probabilitatea ca ținta să fie situată în segmentul a va crește.

Exemplul 2. Probabilitatea ca, într'o fabrică oarecare, produsele să corespundă standardului este de 0,96. Se propune un sistem simplificat de încercări [1] care, pentru produsele corespunzătoare standardului, dă un rezultat pozitiv cu probabilitatea 0,98, iar pentru produsele necorespunzătoare, cu probabilitate de numai

[1] Necesitatea unui control simplificat se întâlnește foarte des în practică. De exemplu, dacă în producția de becuri electrice le-am supune pe toate la o verificare a capacității de ardere timp de cel puțin 1 200 ore, consumatorul ar primi numai becuri arse sau aproape arse. Este deci necesar să se înlocuiască încercarea duratei de ardere cu o altă încercare, de exemplu cu încercarea becului la aprindere.

Scanned with OKEN Scanner

0,05. Care este probabilitatea ca un produs supus de două ori unei încercări simplificate să corespundă standardului?

Aici sistemul complect de ipoteze constă din două evenimente opuse: 1. produsul corespunde standardului, 2. produsul nu corespunde standardului. Probabilităţile acestor ipoteze până la efectuarea încercării sunt egale respectiv cu $P_1=0,96$ şi $P_2=0,04$. In prima ipoteză, ca produsul să reziste la încercare, probabilitatea este egală cu $p_1=0,98$, iar în a doua ipoteză — $p_2=0,05$. După o dublă experienţă, probabilitatea primei ipoteze, în baza formulei (11), este egală cu:

$$\frac{P_1 p_1^2}{P_1 p_1^2 + P_2 p_2^2} = \frac{0,96 \cdot (0,98)^2}{0,96 \cdot (0,98)^2 + 0,04 \cdot (0,05)^2} \approx 0,9999.$$

Vedem că, dacă produsul a rezistat la încercările indicate în problemă, numai într'un caz din 10 000 vom comite o eroare presupunând că produsul este corespunzător standardului. Aceasta desigur satisface în întregime necesităţile practicii.

Exemplul 3. La examinarea unui bolnav se bănueşte că suferă de una din trei boli A_1, A_2, A_3. Probabilităţile lor, în condiţiile date, sunt respectiv :

$$P_1=\frac{1}{2}; \quad P_2=\frac{1}{6}; \quad P_3=\frac{1}{3}.$$

Pentru precizarea diagnosticului, s'a propus o analiză care dă rezultate pozitive cu probabilitatea de 0,1 în cazul bolii A_1, cu probalilitatea de 0,2 în cazul bolii A_2 şi cu probabilitatea 0,9 în cazul bolii A_3. Analiza s'a efecuat de cinci ori şi a dat patru rezultate pozitive şi unul negativ. Se cere găsirea probabilităţii fiecărei boli, după analizele efectuate.

In cazul bolii A_1, probabilitatea rezultatelor analizei este, conform regulii de înmulţire, $p_1=C_5^4(0,1)^4 \cdot 0,9$.

Pentru a doua ipoteză, această probabilitate este $p_2=C_5^4(0,2)^4 \cdot 0,8$ şi pentru a treia este $p_3=C_5^4(0,9)^4 \cdot 0,1$.

Aplicând formula lui Bayes găsim că, în urma analizelor, probabilitatea bolii A_1 este:

$$\frac{P_1 p_1}{P_1 p_1 + P_2 p_2 + P_3 p_3} = \frac{\frac{1}{2} \cdot (0,1)^4 \cdot 0,9}{\frac{1}{2}(0,1)^4 \cdot 0,9 + \frac{1}{6} \cdot (0,2)^4 \cdot 0,8 + \frac{1}{3} \cdot (0,9)^4 \cdot 0,1} \approx 0,002;$$

probabilitatea bolii A_2 —

$$\frac{P_2 p_2}{P_1 p_1 + P_2 p_2 + P_3 p_3} = \frac{\frac{1}{6} \cdot (0,2)^4 \cdot 0,8}{\frac{1}{2} \cdot (0,1)^4 \cdot 0,9 + \frac{1}{6} \cdot (0,2)^4 \cdot 0,8 + \frac{1}{3} \cdot (0,9)^4 \cdot 0,1} \approx 0,010;$$

Scanned with OKEN Scanner

probabilitatea bolii A_3 —

$$\frac{P_3 p_3}{P_1 p_1 + P_2 p_2 + P_3 p_3} = \frac{\frac{1}{3} \cdot (0{,}9)^4 \cdot 0{,}1}{\frac{1}{2} \cdot (0{,}1)^4 \cdot 0{,}9 + \frac{1}{6} \cdot (0{,}2)^4 \cdot 0{,}8 + \frac{1}{3} \cdot (0{,}9)^4 \cdot 0{,}1} \approx 0{,}988.$$

Intrucât aceste trei evenimente A_1, A_2, A_3 formează și după experiență un sistem complect de evenimente, pentru verificarea calculului făcut putem aduna cele trei numere obținute și ne vom convinge că suma lor este egală cu unitatea.

CAPITOLUL V

SCHEMA LUI BERNOULLI

§ 13. Exemple

Exemplul 1. Printre fibrele de bumbac de o anumită calitate, 75% în medie au o lungime mai mică de 45 mm, iar 25% au o lungime mai mare (sau egală) de 45 mm. Să se determine probabilitatea ca, din trei fibre luate la întâmplare, două să fie mai scurte, iar una mai lungă de 45 mm.

Să notăm evenimentul — alegerea unei fibre cu lungimea sub 45 mm — cu A, iar evenimentul — alegerea unei fibre cu lungimea peste 45 mm — prin B; este clar, în acest caz, că:

$$P(A) = \frac{3}{4}; \quad P(B) = \frac{1}{4}.$$

Convenim apoi să notăm cu schema AAB evenimentul compus: primele două fibre alese sunt mai scurte decât 45 mm, iar a treia mai lungă de 45 mm. Este clară semnificația schemelor BBA, ABA, etc. Problema ce se pune este de a calcula probabilitatea evenimentului C, care constă într'aceea că, din trei fibre, două vor fi mai scurte decât 45 mm, iar una mai lungă de 45 mm. Pentru aceasta, evident, trebue să se realizeze una din următoarele scheme:

$$AAB, \quad ABA, \quad BAA. \tag{1}$$

Intrucât oricare două rezultate dintre acestea trei sunt incompatibile între ele, conform regulii de adunare avem:

$$P(C) = P(AAB) + P(ABA) + P(BAA).$$

Scanned with OKEN Scanner

Cei trei termeni din membrul al doilea al egalității sunt egali între ei, deoarece putem considera rezultatele alegerii fibrelor ca evenimente reciproc independente. Probabilitatea fiecăreia din schemele (1), conform regulii de înmulțire a probabilităților pentru evenimente independente, se va scrie ca un produs de trei factori, dintre care doi sunt egali cu $P(A)=3/4$, iar unul cu $P(B)=1/4$. Astfel, probabilitatea fiecăreia dintre cele trei scheme (1) este egală cu:

$$\left(\frac{3}{4}\right)^2 \cdot \frac{1}{4}=\frac{9}{64},$$

deci

$$P(C)=3 \cdot \frac{9}{64}=\frac{27}{64},$$

ceeace rezolvă problema.

Exemplul 2. In urma unor observații, care au durat mai multe decenii, s'a constatat că la fiecare mie de noi născuți se nasc în medie 515 băieți și 485 fete. Intr'o familie oarecare sunt 6 copii. Să se determine probabilitatea ca printre aceștia să fie cel mult două fete.

Pentru producerea evenimentului a cărui probabilitate o căutăm, trebue ca în această familie să existe fie 0, fie 1, fie 2 fete. Probabilitățile acestor evenimente particulare le vom nota respectiv P_0, P_1, P_2. Evident că, în conformitate cu regula de adunare, probabilitatea căutată este:

$$P=P_0+P_1+P_2. \qquad (2)$$

Pentru fiecare copil, probabilitatea ca acesta să fie băiat este egală cu 0,515; deci, probabilitatea ca să fie fată este egală cu 0,485.

Cel mai simplu se determină P_0; aceasta este probabilitatea ca toți copiii din familie să fie băieți. Intrucât putem considera evenimentul — nașterea unui copil de un sex sau altul — ca fiind un eveniment independent de nașterea celorlalți copii, conform regulii de înmulțire, probabilitatea ca toți șase copii să fie băieți este egală cu produsul a 6 factori egali cu 0,515, adică:

$$P_0=(0{,}515)^6 \approx 0{,}018.$$

Să calculăm acum probabilitatea P_1, adică aceea ca printre cei 6 copii din familie, unul să fie fată, iar ceilalți cinci, băieți. Acest eveniment se poate produce în șase moduri deosebite, după al câtelea din copii, în ordinea nașterii lor, va fi fată (primul, al doilea, etc.). Să examinăm acum una din variantele evenimentului nostru, de exemplu, aceea că al patrulea din copiii născuți va fi fată. Probabilitatea acestei variante, conform regulii de în-

mulțire, este egală cu produsul a șase factori, dintre care cinci sunt egali cu 0,515, iar al șaselea (care ocupă locul al patrulea) egal cu 0,485, adică, această probabilitate este egală cu $(0,515)^5 \cdot 0,485$. Probabilitatea este aceeași și pentru oricare alta din celelalte cinci variante posibile ale evenimentului care ne interesează; deaceea probabilitatea P_1 a acestui eveniment, conform regulii de adunare este egală cu suma a șase numere egale fiecare cu $(0,515)^5 \cdot 0,485$, adică:

$$P_1 = 6 \cdot (0,515)^5 \cdot 0,485 \approx 0,105.$$

Să trecem acum la calculul lui P_2 (probabilitatea ca doi din copii să fie fete, iar patru — băieți). Ca și mai sus, vedem imediat că acest eveniment admite o serie întreagă de variante (una din variante va fi, de exemplu, următoarea: al doilea și al cincilea copil, în ordinea nașterii lor, vor fi fete, iar ceilalți copii — băieți). Probabilitatea fiecărei variante este, conform regulii de înmulțire, egală cu $(0,515)^4 \cdot (0,485)^2$, deci P_2, conform regulii de adunare, este egală cu $(0,515)^4 \cdot (0,485)^2$, înmulțit cu numărul variantelor posibile de tipul considerat; problema se reduce deci la determinarea acestui ultim număr.

Fiecare variantă este caracterizată prin aceea, că, din șase copii doi sunt fete, iar ceilalți băieți; numărul diferitelor variante este deci egal cu numărul diferitelor moduri de a alege doi copii din șase. Numărul unor astfel de alegeri este egal cu numărul combinărilor de șase obiecte luate câte două, adică $C_6^2 = \frac{6 \cdot 5}{2 \cdot 1} = 15$.

Prin urmare,

$$P_2 = C_6^2 \ (0,515)^4 \cdot (0,485)^2 = 15 \cdot (0,515)^4 \cdot (0,485)^2 \approx 0,247.$$

Adunând rezultatele obținute rezultă:

$$P = P_0 + P_1 + P_2 \approx 0,018 + 0,105 + 0,247 = 0,370.$$

Astfel, ceva mai puțin frecvent ca în patru cazuri din zece, în asemenea familii cu mai mulți copii, cel mult a treia parte din ei vor fi fete și cel puțin două treimi — băieți.

§ 14. Formulele lui Bernoulli

În paragraful precedent, am văzut, într'un număr de exemple, ce înseamnă schema *repetării probelor* când în fiecare din ele se poate produce un eveniment oarecare A. Vom da cuvântului „probă" o semnificație foarte largă și variată. Astfel, dacă efectuăm o tragere într'o țintă vom înțelege prin probă fiecare foc tras în parte. Dacă facem încercarea becurilor electrice pentru durata

Scanned with OKEN Scanner

arderii, prin probă vom înțelege verificarea fiecăruia din becuri. Dacă cercetăm distribuția noilor născuți, pe sex, greutate sau lungime, vom înțelege prin probă examinarea fiecărui copil. In general, prin probă vom înțelege în cele ce urmează realizarea unor anumite condiții, în care se poate produce un anumit eveniment, care ne interesează.

Am trecut acum la studiul uneia dintre schemele principale ale calculului probabilităților, care, în afară de importanța aplicării ei în diferite domenii ale științei, are o mare însemnătate în însuși calculul probabilităților, ca știință matematică. In această schemă se studiază un șir de probe reciproc independente, adică astfel de probe încât probabilitatea unui rezultat sau altuia al fiecăreia dintre ele nu depinde de rezultatele care s'au produs sau se vor produce în celelalte probe. In fiecare dintre aceste probe, poate să se producă (sau să nu se producă) un eveniment oarecare A, cu probabilitatea p, care nu depinde de numărul de ordine al probei. Schema descrisă a căpătat denumirea de *schema lui Bernoulli*, întrucât studiul ei sistematic a fost făcut de cunoscutul savant elvețian Jacob Bernoulli, care a trăit spre sfârșitul secolului al XVII-lea.

Am mai întâlnit, în exemple, schema lui Bernoulli; pentru a ne convinge, este suficient să ne aducem aminte de exemplele din paragraful precedent. Acum, însă, să rezolvăm următoarea problemă generală, pentru care toate exemplele examinate până acum, în acest capitol, sunt cazuri particulare.

Problemă. In anumite condiții, probabilitatea apariției evenimentului A în fiecare probă este egală cu p; să se determine probabilitatea ca, într'o serie de n probe independente, evenimentul A să se producă de k ori și să nu se producă de $n—k$ ori.

Evenimentul a cărui probabilitate o căutăm se descompune într'un număr de variante; pentru a obține o anumită variantă, trebue să luăm în mod arbitrar k probe din seria dată și să presupunem că evenimentul A s'a produs chiar în aceste k probe, iar în celelalte $n—k$ probe el nu s'a produs. Astfel, fiecare din aceste variante necesită apariția a n rezultate determinate, în care să avem de k ori producerea evenimentului A și de $n—k$ ori — neproducerea lui; obținem, conform regulii de înmulțire, că probabilitatea fiecărei variante determinate este:

$$p^k(1-p)^{n-k}.$$

Numărul diferitelor variante posibile este egal cu numărul diferitelor grupe de k probe ce se pot forma din n probe, adică este egal cu C_n^k. Aplicând regula de adunare a probabilităților și formula cunoscută pentru combinări

$$C_n^k = \frac{n(n-1)\dots[n-(k-1)]}{k(k-1)\dots 2\cdot 1},$$

găsim că probabilitatea ca evenimentul A să se producă de k ori în n probe independente este:

$$P_n(k) = \frac{n(n-1)\dots[n-(k-1)]}{k(k-1)\dots 2-1} p^k (1-p)^{n-k} \qquad (3)$$

ceeace rezolvă problema pusă. Se întâmplă deseori să fie mai comod să se pună expresia C^k sub o formă puțin schimbată; înmulțind numărătorul și numitorul cu produsul $(n-k)\ [n-(k+1)]\dots 2:1$, obținem:

$$C_n^k = \frac{n(n-1)\dots 2\cdot 1}{k(k-1)\dots 2\cdot 1\cdot (n-k)\,[n-(k+1)]\dots 2\cdot 1},$$

sau însemnând pentru prescurtare prin $m!$ produsul tuturor numerelor întregi dela 1 până la m incluziv,

$$C_n^k = \frac{n!}{k!\,(n-k)!};$$

pentru $P_n(k)$, aceasta dă:

$$P_n(k) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}. \qquad (4)$$

Formulele (3) și (4) se numesc, deobicei, *formulele lui Bernoulli*. Pentru valori mari ale lui n și k, calculul lui $P_n(k)$ cu aceste formule este greoi, întrucât factorialii $n!$, $k!$, $(n-k)!$ — sunt numere foarte mari și necesită un calcul lung. Deaceea, în astfel de calcule sunt folosite pe o scară largă atât tabelele special întocmite pentru factoriali cât și unele formule aproximative.

Exemplu. Probabilitatea ca la o întreprindere oarecare consumul de apă să fie normal (să nu depășească un număr determinat de litri în 24 de ore) este egală cu $\frac{3}{4}$. Să se găsească probabilitatea ca în șase zile consecutive consumul de apă să fie normal timp de 0, 1, 2, 3, 4, 5, 6 zile.

Notând cu $P_6(k)$ probabilitatea ca timp de k zile din șase, consumul de apă să fie normal, obținem conform formulei (3) $\left(\text{în care punem } p=\frac{3}{4}\right)$:

$$P_6(6) = \left(\frac{3}{4}\right)^6 = \frac{3^6}{4^6},$$

$$P_6(5) = 6\cdot\left(\frac{3}{4}\right)^5\cdot\frac{1}{4} = \frac{6\cdot 3^5}{4^6},$$

$$P_6(4) = C_6^4\cdot\left(\frac{3}{4}\right)^4\cdot\left(\frac{1}{4}\right)^2 = C_6^2\frac{3^4}{4^6} = \frac{6\cdot 5}{2\cdot 1}\cdot\frac{3^4}{4^6} = \frac{15\cdot 3^4}{4^6},$$

Scanned with OKEN Scanner

$$P_6(3) = C_6^3 \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^3 = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} \cdot \frac{3^3}{4^6} = \frac{20 \cdot 3^3}{4^6},$$

$$P_6(2) = \frac{6 \cdot 5}{2 \cdot 1} \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^4 = \frac{15 \cdot 3^2}{4^6},$$

$$P_6(1) = 6 \cdot \frac{3}{4} \cdot \left(\frac{1}{4}\right)^5 = \frac{6 \cdot 3}{4^6};$$

în sfârșit, este evident că $P_6(0)$ (probabilitatea de a depăși consumul în fiecare din cele șase zile) este $\frac{1}{4^6}$. Toate cele șase probabilități sunt exprimate prin fracții având același numitor $4^6 = 4\,096$; desigur că ne folosim de aceasta pentru simplificarea calculelor. Rezultatele calculelor sunt:

$P_6(6) \approx 0{,}18$; $P_6(5) \approx 0{,}36$; $P_6(4) \approx 0{,}30$:
$P_6(3) \approx 0{,}13$; $P_6(2) \approx 0{,}03$; $P_6(1) \approx P_6(0) \approx 0$.

Se vede că cel mai probabil este să avem depășiri de consum de apă într'o zi sau în două zile din cele șase și că probabilitatea depășirii consumului în timp de 5 sau 6 zile $[P_6(1) + P_6(0)]$ este practic egală cu zero.

§ 15. Numărul cel mai probabil de apariții ale unui eveniment

Exemplul pe care l-am examinat arată că probabilitatea supraconsumului de apă, timp de k zile, crește la început pe măsură ce crește numărul k iar apoi, atingând valoarea sa maximă, începe să scadă; aceasta se vede în modul cel mai clar, dacă reprezentăm geometric variația probabilității $P_6(k)$, în funcție de creșterea numărului k, sub forma diagramei arătate în fig. 4

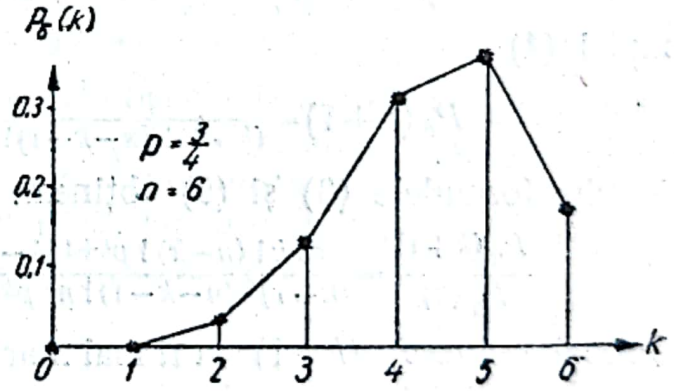

Fig. 4

Diagrama variației mărimii $P_n(k)$ în funcție de k, dă o imagine cu atât mai vie cu cât n este mai mare; astfel, pentru $n=15$ și $p=\frac{1}{2}$, diagrama are forma arătată în fig. 5.

In practică, se cere uneori să cunoaștem numărul cel mai probabil de produceri ale unui eveniment, adică numărul k, pentru care probabilitatea $P_n(k)$ este maximă (presupunând, bineînțeles, că p și n sunt date).

Scanned with OKEN Scanner

Formulele lui Bernoulli ne permit să găsim în toate cazurile o soluție simplă a acestei probleme; cu aceasta ne vom ocupa acum.

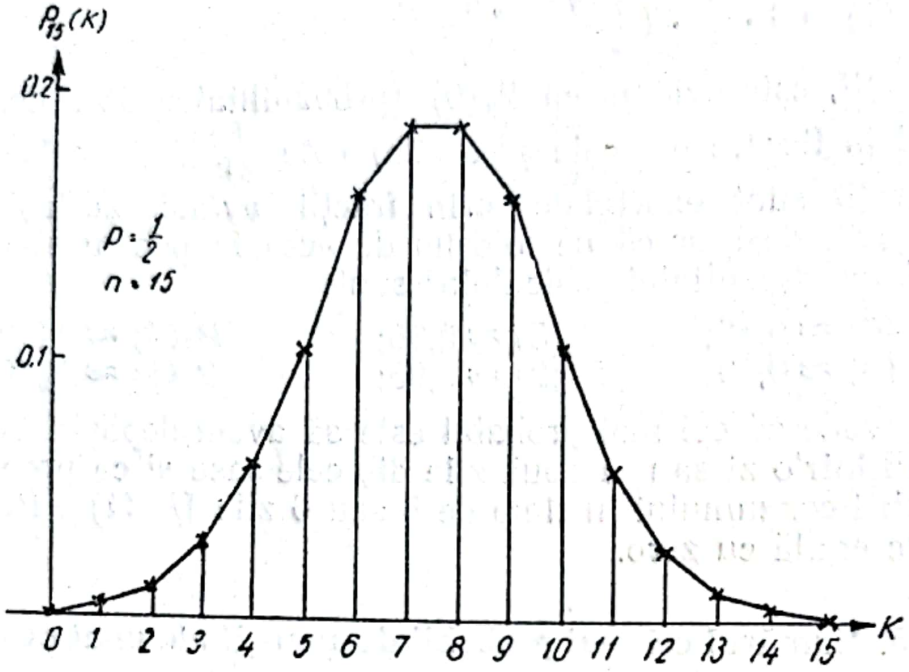

Fig. 5

Să calculăm întâi valoarea raportului $\frac{P_n(k+1)}{P_n(k)}$. Conform formulei (4)

$$P_n(k+1)=\frac{n!}{(k+1)!\,(n-k-1)!}\,p^{k+1}\,(1-p)^{n-k-1}, \qquad (5)$$

iar din formulele (3) și (5) obținem:

$$\frac{P_n(k+1)}{P_n(k)}=\frac{n!\,k!\,(n-k)!\,p^{k+1}\,(1-p)^{n-k-1}}{(k+1)!\,(n-k-1)!\,n!\,p^k\,(1-p)^{n-k}}=\frac{n-k}{k+1}\,\frac{p}{1-p}.$$

Probabilitatea $P_n(k+1)$ va fi mai mare, egală sau mai mică decât probabilitatea $P_n(k)$, după cum raportul $\frac{P_n(k+1)}{P_n(k)}$ este mai mare, egal sau mai mic decât unu, iar aceasta, precum se vede, se reduce la a determina care dintre cele trei relații

$$\frac{n-k}{k+1}\cdot\frac{p}{1-p}>1, \qquad \frac{n-k}{k+1}\cdot\frac{p}{1-p}=1, \qquad \frac{n-k}{k+1}\cdot\frac{p}{1-p}<1 \qquad (6)$$

este adevărată. Dacă vrem, de exemplu, să găsim valorile lui k pentru care are loc inegalitatea $P_n(k+1)>P_n(k)$, trebue să de-

terminăm valorile lui k, pentru care este valabilă inegalitatea

$$\frac{n-k}{k+1} \cdot \frac{p}{1-p} > 1,$$

sau

$$(n-k)p > (k+1)(1-p),$$

de unde

$$np-(1-p) > k;$$

astfel, cât timp k, crescând, nu va ajunge la valoarea $np-(1-p)$, vom avea mereu $P_n(k+1) > P_n(k)$, adică cu cât numărul k crește, cu atât crește și probabilitatea $P_n(k)$. Astfel, în schema căreia îi corespunde diagrama din fig. 5, $p=\frac{1}{2}$, $n=15$, $np-(1-p)=7$; deci câtă vreme $k<7$, adică pentru valorile lui k cuprinse între 0 și șase incluziv, $P_n(k+1) > P_n(k)$. Diagrama confirmă acest lucru.

In mod analog, pornind dela celelalte două relații (6) rezultă că:

$$P_n(k+1) = P_n(k), \quad \text{dacă} \quad k = np-(1-p),$$

și

$$P_n(k+1) < P_n(k), \quad \text{dacă} \quad k > np-(1-p);$$

deci, îndată ce numărul k, crescând, va depăși pe $np-(1-p)$, probabilitatea $P_n(k)$ va începe să scadă până la $P_n(n)$.

Această concluzie ne arată, înainte de toate, că comportarea mărimii $P_n(k)$ observată în exemplele de mai sus, în funcție de creșterea numărului k (întâi creștere, apoi descreștere) este o lege generală valabilă în toate cazurile.

Mai mult decât atât, această concluzie ne permite deasemenea să rezolvăm imediat problema pe care ne-am propus-o și anume determinarea valorii celei mai probabile a numărului k. Să însemnăm valoarea cea mai probabilă prin k_0. Avem:

$$P_n(k_0+1) \leq P_n(k_0),$$

de unde, conform celor de mai sus, rezultă:

$$k \geq np-(1-p);$$

pe de altă parte

$$P_n(k_0-1) \leq P_n(k_0),$$

iar pentru aceasta, conform rezultatelor precedente, trebue să aibă loc inegalitatea:

$$k_0-1 \leq np-(1-p)$$

sau

$$k_0 \leq np-(1-p)+1 = np+p.$$

Scanned with OKEN Scanner

Astfel, valoarea cea mai probabilă k_0 a numărului k, trebue să satisfacă dubla inegalitate:

$$np-(1-p) \leq k_0 \leq np+p; \qquad (7)$$

intervalul dintre $np-(1-p)$ și $np+p$, care trebue să conțină numărul k_0, are lungimea 1, ceeace se vede printr'o simplă scădere; deaceea, dacă una dintre extremitățile acestui interval, de exemplu numărul $np-(1-p)$, nu este un număr întreg, între aceste extremități se va găsi în mod obligatoriu un număr întreg și unul singur, deci k_0 este determinat în mod unic. Trebue să considerăm acest caz ca fiind cazul normal: $p<1$; deaceea, numai în cazuri excepționale, mărimea $np-(1-p)$ va fi un număr întreg. In acest caz excepțional, inegalitățile (7) dau pentru numărul k_0 două valori: $np-(1-p)$ și $np+p$, care diferă între ele cu o unitate; aceste două valori vor fi valorile cele mai probabile; probabilitățile lor sunt egale între ele și sunt mai mari decât probabilitățile corespunzătoare tuturor celorlalte valori ale numărului k. Tocmai acest caz excepțional are loc, de exemplu, pe schema reprezentată în diagrama din fig. 5, unde $n=15$, $p=\frac{1}{2}$, deci $np-(1-p)=7$, $np+p=8$; numerele 7 și 8 sunt valorile cele mai probabile ale numărului k pentru ca evenimentul să se producă; probabilitățile lor sunt egale între ele, fiecare dintre ele fiind egală aproximativ cu 0,196 (toate acestea se pot vedea din diagramă).

Exemplul 1. In urma observațiilor făcute timp de mai mulți ani, într'o localitate oarecare s'a stabilit că probabilitatea ca în ziua de 1 Iulie să plouă este egală cu 4/17. Să se determine numărul cel mai probabil de zile de 1 Iulie ploioase din ultimii 50 de ani. In acest caz $n=50$, $p=4/17$

$$np-(1-p)=50\cdot\frac{4}{17}-\frac{13}{17}=11;$$

am obținut un rezultat cu un număr întreg; deci avem de a face cu un caz excepțional; valorile cele mai probabile ale numărului de zile ploioase vor fi numerele 11 și 12, probabilitățile lor fiind egale între ele.

Exemplul 2. Tunul Nr. 1 poate să tragă într'un interval de timp dat 60 de obuze cu o probabilitate de nimerire de 0,7 pentru fiecare dintre ele; iar tunul Nr. 2, în același interval de timp, — 50 de obuze cu o probabilitate de nimerire de 0,8; pentru care din aceste două tunuri numărul cel mai probabil de nimeriri la țintă este mai mare?

Pentru tunul Nr. 1:

$$n=60, \quad p=0,7, \quad np-(1-p)=41,7, \quad k_0=42.$$

Scanned with OKEN Scanner

Pentru tunul Nr. 2:

$$n=50, \quad p=0{,}8, \quad np-(1-p)=39{,}8, \quad k_0=40.$$

Vedem că numărul cel mai probabil de nimeriri este mai mare la tunul Nr. 1 decât la tunul Nr. 2.

Se întâlnesc adesea în practică situații în care numărul n este foarte mare (tragere de masă, fabricare de produse în masă, etc.). In acest caz, produsul np va fi deasemenea un număr foarte mare (dacă probabilitatea p nu este extrem de mică). Deoarece în expresiile $np-(1-p)$ și $np+p$, între care este cuprins numărul cel mai probabil de apariții ale evenimentului, termenii ai doilea (adică p și $1-p$) sunt mai mici decât unu, amândouă aceste numere, deci și numărul cel mai probabil de apariții ale evenimentului cuprins între ele, sunt apropiate de np. Astfel, trăgând 1 000 focuri, cu o probabilitate de nimerire de 0,74 pentru fiecare dintre ele, putem lua $1\,000 \cdot 0{,}74=740$ ca valoarea cea mai probabilă pentru numărul de nimeriri.

Putem da acestui rezultat o formă mai precisă. Dacă k_0 reprezintă numărul cel mai probabil de produceri ale unui eveniment în n probe, $\frac{k_0}{n}$ este cea mai probabilă „fracțiune" de produceri ale evenimentului în aceste n probe. Inegalitățile (7) dau:

$$p-\frac{1-p}{n} \leq \frac{k_0}{n} \leq p+\frac{p}{n}.$$

Să presupunem că, menținând constantă probabilitatea p de producere a evenimentului într'o probă izolată, vom mări din ce în ce mai mult numărul n al probelor (atunci, bineînțeles, va crește și numărul cel mai probabil k_0). Fracțiile $\frac{1-p}{n}$ și $\frac{p}{n}$ care se află în membrii extremi ai dublei inegalități precedente vor deveni din ce în ce mai mici; deci, pentru o valoare foarte mare a lui n, putem neglija aceste fracții, adică putem considera membrii extremi ai inegalității, deci și fracțiunea $\frac{k_0}{n}$ cuprinsă între ei, egali cu p. Astfel, *cea mai probabilă fracțiune de producere ale unui eveniment, la un număr mare de probe, este practic egală cu probabilitatea producerii evenimentului într'o singură probă.* Astfel, dacă la o tragere probabilitatea nimeririi este de 0,84 pentru un singur foc, într'un număr mare de focuri trase ne putem aștepta cu cea mai mare probabilitate la circa 84% nimeriri.

Aceasta nu înseamnă desigur că probabilitatea de a obține exact 84% nimeriri este foarte mare; dimpotrivă, însăși această „probabilitate maximă" va fi foarte mică când numărul de focuri

Scanned with OKEN Scanner

este mare (astfel, am văzut în schema fig. 5 că probabilitatea maximă era egală cu 0,196, fiind vorba numai de 15 probe: la un număr mai mare de probe, ea este simţitor mai mică). Această probabilitate este „maximă" numai relativ: probabilitatea de a obţine 84% nimeriri este mai mare decât probabilitatea de a obţine 83% nimeriri.

Pe de altă parte se înţelege că la o tragere în serii lungi, probabilitatea unui număr sau altul de nimeriri nu poate prezenta un interes deosebit. Dacă, de exemplu, lansăm 200 de mine nici nu este necesar să se calculeze probabilitatea ca exact 137 din ele să nimerească ţinta, pentru că, din punct de vedere practic, este indiferent dacă ţinta va fi nimerită de 137, 136, 138 sau chiar 140 de ori. Dimpotrivă, problema găsirii, de exemplu, a probabilităţii ca ţinta să fie nimerită de mai mult de 100 din cele 200 de mine aruncate, sau ca numărul de nimeriri să fie cuprins între 100 şi 125, sau ca acesta să fie mai mic decât 50, etc., prezintă un indiscutabil interes practic. Cum să cercetăm acest fel de probabilităţi? Să presupunem, de exemplu, că vrem să găsim probabilitatea ca numărul de nimeriri să fie cuprins între 100 şi 120 (inclusiv 120); mai precis, vom căuta probabilitatea inegalităţilor

$$100 < k \leq 120,$$

unde k este numărul de nimeriri. Pentru ca aceste inegalităţi să aibă loc, trebue ca numărul k să fie egal cu unul dintre cele 20 de numere, 101, 102, ..., 120; conform regulii de adunare, această probabilitate este:

$$P(100 < k \leq 120) = P_{200}(101) + P_{200}(102) + \ldots + P_{200}(120);$$

Pentru calculul direct al acestei sume trebue să calculăm în prealabil 20 probabilităţi diferite de tipul $P_n(k)$, cu ajutorul formulei (3); pentru numere atât de mari, astfel de calcule prezintă greutăţi de neînvins; deaceea, în practică asemenea sume nu sunt calculate niciodată în mod direct. Există în acest scop formule aproximative şi tabele care uşurează calculul. Întocmirea acestor formule şi tabele este bazată pe procedeele complicate ale analizei matematice, de care nu ne vom ocupa aici; cu toate acestea se pot obţine date şi despre probabilităţi de tipul $P(100 < k \leq 120)$ numai prin simple raţionamente, care în multe cazuri duc la o rezolvare completă a problemelor puse. Despre aceasta, vom vorbi în capitolul următor.

Scanned with OKEN Scanner

CAPITOLUL VI

TEOREMA LUI BERNOULLI

§ 16. Conținutul teoremei lui Bernoulli

Să examinăm încă odată diagrama din fig. 5 (pag. 42), unde probabilitățile diferitelor valori ale numărului k de produceri ale evenimentului considerat, adică numerele $P_{15}(k)$, sunt reprezentate prin linii verticale. Probabilitatea corespunzătoare unui interval oarecare de valori ale lui k, adică probabilitatea ca evenimentul care ne interesează să se producă de un număr de ori egal cu un număr oarecare din acest interval, este egală, conform regulii de adunare, cu suma probabilităților tuturor numerelor cuprinse în acest interval, adică cu suma lungimilor tuturor liniilor verticale, situate în acest interval. Figura ne arată în mod clar că această sumă diferă foarte mult dela un interval la altul de aceeași lungime. Astfel, intervalele $2 \leq k < 5$ și $7 \leq k < 10$ au aceeași lungime. Probabilitatea fiecăruia se exprimă prin suma lungimilor a trei linii verticale și vedem că în al doilea interval această sumă este mult mai mare decât în primul. Știm că diagramele probabilităților $P_n(k)$ au în esență aceeași formă pentru orice număr n ca și diagrama din fig. 5, adică știm că mărimea $P_n(k)$, pe măsură ce k crește, crește și ea iar apoi, după ce trece prin valoarea ei maximă, începe să descrească; deaceea, este clar că, pentru două intervale de valori ale numărului k, având aceeași lungime, întotdeauna o probabilitate mai mare va corespunde intervalului situat mai aproape de valoarea cea mai probabilă k_0. In particular, intervalul cu centrul în k_0 va fi acela căruia îi corespunde o probabilitate mai mare decât oricărui alt interval de aceeași lungime.

Vom vedea însă, în legătură cu aceasta, că se poate spune mult mai mult. Numărul total de valori posibile ale numărului k de produceri ale unui eveniment în n probe este egal cu $n+1\,(0 \leq k \leq n)$.

Să considerăm segmentul cu centrul în k_0 și cuprinzând o mică parte, de exemplu o sutime, din toate valorile posibile ale numărului k: se constată, în acest caz, că dacă numărul total n de probe este foarte mare, acestui segment îi corespunde o probabilitate considerabilă, iar tuturor celorlalte valori ale numărului k, luate împreună, le corespunde o probabilitate extrem de mică. Astfel, cu toate că segmentul considerat este extrem de mic în comparație cu n (pe figură acest segment ocupă numai a suta parte din lungimea diagramei), totuși suma liniilor verticale

Scanned with OKEN Scanner

situate deasupra lui este mult mai mare decât suma tuturor celorlalte linii verticale. Aceasta se datorește faptului că liniile părții centrale a diagramei sunt mult mai mari decât liniile situate la extremitățile ei. Astfel, pentru n mare, diagrama $P_n(k)$ are aproximativ forma din fig. 6.

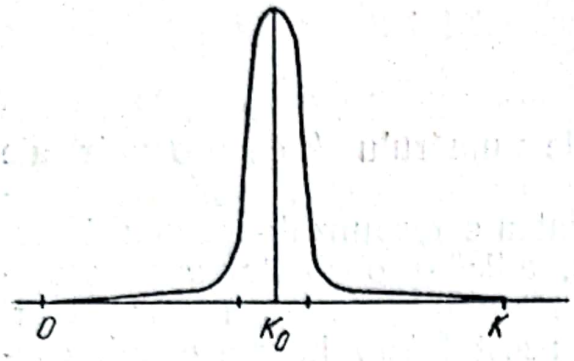

Fig. 6

Din punct de vedere practic aceasta înseamnă, evident, următoarele: *dacă efectuăm un număr mare n de probe, ne putem aștepta, cu o probabilitate apropiată de unitate, ca numărul k de produceri ale evenimentului A să fie foarte apropiat de valoarea sa cea mai probabilă, diferind de ea doar cu o parte neînsemnată din numărul total n al probelor făcute.*

Această propoziție, cunoscută sub numele de *teorema lui Bernoulli*, care a fost enunțată la începutul secolului al XVIII-lea, reprezintă una din legile cele mai importante ale teoriei probabilităților; până la mijlocul secolului trecut, toate demonstrațiile acestei teoreme cereau un aparat matematic complicat. Marele matematician rus P. L. Cebâșev a fost primul care a dat o demonstrație foarte simplă și scurtă a acestei legi; dăm aici remarcabila demonstrație a lui Cebâșev.

§ 17. Demonstrația teoremei lui Bernoulli

Știm că, dacă numărul n de probe este foarte mare, numărul cel mai probabil k_0 de produceri ale evenimentului A diferă foarte puțin de valoarea np, unde p, ca întotdeauna, reprezintă probabilitatea evenimentului A, într'o singură probă. Deaceea este suficient să demonstrăm că, la un număr mare de probe, probabilitatea ca numărul k de produceri ale evenimentului A să difere foarte puțin de np — cu mai puțin decât o fracțiune suficient de mică din n (cel mult de exemplu, cu 0,01 n, sau cu 0,001 n, sau în general cu cel mult εn, unde ε este un număr oricât de mic) este foarte mare; cu alte cuvinte, trebue să demonstrăm că pro-

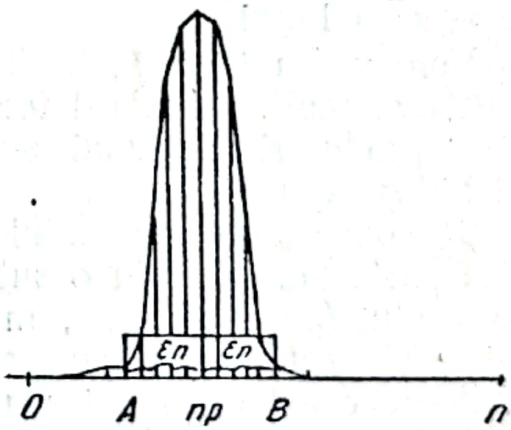

Fig. 7

babilitatea

$$P(|k-np|>\varepsilon n) \qquad (1)$$

va fi oricât de mică, pentru valori suficient de mari ale lui n.

Pentru a ne convinge de acest lucru, să observăm că, în conformitate cu regula de adunare, probabilitatea (1) este egală cu suma probabilităţilor $P_n(k)$ pentru toate valorile lui k, care diferă de np cu mai mult decât εn; pe diagrama obişnuită (fig. 7), această sumă este reprezentată de suma lungimilor tuturor liniilor verticale situate în afara segmentului AB, — atât în dreapta cât şi în stânga lui. Întrucât suma tuturor liniilor verticale (ca fiind suma probabilităţilor unui sistem complect de evenimente) este egală cu unitatea, aceasta înseamnă că partea covârşitoare a acestei sume, apropiindu-se de unitate, corespunde segmentului AB şi numai o parte extrem de mică a sumei corespunde intervalelor în afara acestui segment.

Astfel,

$$P(|k-np|>\varepsilon n)=\sum_{|k-np|>\varepsilon n} P_n(k). \qquad (2)$$

Să recurgem acum la raţionamentul lui Cebâşev. Întrucât pentru fiecare din termenii sumei scrise

$$\left|\frac{k-np}{\varepsilon n}\right|>1,$$

şi deci

$$\left(\frac{k-np}{\varepsilon n}\right)^2>1,$$

putem mări această sumă dacă înlocuim fiecare termen al ei $P_n(k)$ prin expresia

$$\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k);$$

deaceea

$$P(|k-np|>\varepsilon n)<\sum_{|k-np|>\varepsilon n}\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k)=\frac{1}{\varepsilon^2 n^2}\sum_{|k-np|>\varepsilon n}(k-np)^2 P_n(k);$$

de unde rezultă că ultima sumă se va mări şi mai mult dacă îi vom adăuga noi termeni, făcându-l pe k să parcurgă nu numai intervalele până la $np-\varepsilon n$ şi dela $np+\varepsilon n$, ci toată mulţimea valorilor posibile, adică toate numerele dela 0 până la n incluziv; rezultă deci cu atât mai mult

$$P(|k-np|>\varepsilon n)<\frac{1}{\varepsilon^2 n^2}\sum_{k=0}^{n}(k-np)^2 P_n(k). \qquad (3)$$

Această ultimă sumă prezintă un avantaj faţă de toate sumele precedente, prin faptul că poate fi calculată cu precizie; metoda

lui Cebâșev constă în înlocuirea unei sume greu de calculat (2) prin suma (3) care poate fi calculată exact.

Să trecem acum la efectuarea acestui calcul: oricât de greu ni s'ar părea, acest lucru nu prezintă decât greutăți de ordin tehnic, cărora le poate face față oricine, cu cunoștințe de algebră. Remarcabila idee a lui Cebâșev a și fost din plin folosită la noi, deoarece ea constă chiar în trecerea dela egalitatea (2) la inegalitatea (3).

Obținem cu ușurință în primul rând

$$\sum_{k=0}^{n}(k-np)^2 P_n(k)=\sum_{k=0}^{n}k^2 P_n(k)-2np\sum_{k=0}^{n}k P_n(k)+n^2p^2\sum_{k=0}^{n}P_n(k). \quad (4)$$

Din cele trei sume din membrul al doilea, ultima sumă este egală cu unu, ca fiind suma probabilităților unui sistem complect de evenimente. Ne rămâne, deci, de calculat sumele

$$\sum_{k=0}^{n}k P_n(k) \quad \text{și} \quad \sum_{k=0}^{n}k^2 P_n(k);$$

unde în ambele sume termenii corespunzători lui $k=0$ sunt egali cu 0, așa că putem începe calculul sumei dela $k=1$.

1. Pentru calculul celor două sume îl vom exprima pe $P_n(k)$ cu ajutorul formulei (4) din capitolul V (pag. 40). Obținem:

$$\sum_{k=1}^{n}k P_n(k)=\sum_{k=1}^{n}\frac{kn!}{k!(n-k)!}p^k(1-p)^{n-k};$$

dar întrucât e evident că $n!=n(n-1)!$ și $k!=k(k-1)!$ rezultă că

$$\sum_{k=1}^{n}k P_n(k)=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}(1-p)^{(n-1)-(k-1)},$$

sau, punând în suma din membrul doi $k-1=l$ și observând că l variază dela 0 la $n-1$, când k variază dela 1 la n,

$$\sum_{k=1}^{n}k P_n(k)=np\sum_{l=0}^{n-1}\frac{(n-1)!}{l!(n-1-l)!}p^l(1-p)^{n-1-l}=np\sum_{l=0}^{n-1}P_{n-1}(l).$$

Ultima sumă $\sum_{l=0}^{n-1}P_{n-1}(l)$ este evident egală cu unu, deoarece este suma probabilităților unui sistem complect de evenimente și anume al tuturor numerelor posibile l de producere a evenimentului, în $n-1$ probe. Astfel obținem pentru suma $\sum_{k=0}^{n}k P_n(k)$ o expresie foarte simplă

$$\sum_{k=0}^{n}k P_n(k)=np. \quad (5)$$

2. Pentru calculul sumei a doua vom căuta mai întâi valoarea $\sum_{k=1}^{n} k(k-1)P_n(k)$; întrucât termenul care corespunde lui $k=1$ este egal cu 0, putem începe calculul sumei dela $k=2$. Observând că $n!=n(n-1)(n-2)!$ și că $k!=k(k-1)(k-2)!$ constatăm imediat, notând ca mai sus $k-2=m$:

$$\sum_{k=1}^{n} k(k-1)P_n(k)=\sum_{k=2}^{n} k(k-1)P_n(k)=\sum_{k=2}^{n}\frac{k(k-1)n!}{k!(n-k)!}p^k(1-p)^{n-k}=$$

$$=n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)![(n-2)-(k-2)]!}p^{k-2}(1-p)^{(n-2)-(k-2)}=$$

$$=n(n-1)p^2\sum_{m=0}^{n-2}\frac{(n-2)!}{m!(n-2-m)!}p^m(1-p)^{n-2-m}=$$

$$=n(n-1)p^2\sum_{m=0}^{n-2}P_{n-2}(m)=n(n-1)p^2, \qquad (6)$$

deoarece ultima sumă este din nou egală cu unitatea, ca fiind suma probabilităţilor unui anumit sistem complect de evenimente şi anume a tuturor numerelor posibile de produceri ale evenimentului, în $n-2$ probe.

In sfârşit, formulele (5) şi (6) dau:

$$\sum_{k=1}^{n} k^2P_n(k)=\sum_{k=1}^{n} k(k-1)P_n(k)+\sum_{k=1}^{n} kP_n(k)=n(n-1)p^2+np=$$

$$=n^2p^2+np(1-p). \qquad (7)$$

Acum ambele sume de care avem nevoie sunt calculate. Introducând rezultatele (5) şi (7) în relaţia (4), obţinem în cele din urmă:

$$\sum_{k=0}^{n}(k-np)^2P_n(k)=n^2p^2+np(1-p)-2np\cdot np+n^2p^2=np(1-p);$$

şi introducând expresia extrem de simplă pe care am obţinut-o în inegalitatea (3), obţinem:

$$P(|k-np|>\varepsilon n)<\frac{np(1-p)}{\varepsilon^2n^2}=\frac{p(1-p)}{\varepsilon^2n}. \qquad (8)$$

Această inegalitate demonstrează tot ce ne este necesar. In adevăr, am putut lua un număr ε oricât de mic, dar odată ales nu-l mai schimbăm; în ceeace priveşte numărul n de experienţe, acesta, în sensul afirmaţiilor noastre, poate fi oricât de mare; deaceea fracţia $\frac{p(1-p)}{\varepsilon^2n}$ poate fi presupusă oricât de mică, deoarece

cu cât n crește, numitorul ei devine mai mare, numărătorul rămânând același.

Fie, de exemplu, $p=0{,}75$, astfel încât $1-p=0{,}25$ și

$$p(1-p)=0{,}1875<0{,}2;$$

să alegem $\varepsilon=0{,}01$; atunci din inegalitatea (8) rezultă

$$P\left(\left|k-\frac{3}{4}n\right|>\frac{1}{100}n\right)<\frac{0{,}2}{0{,}0001\cdot n}=\frac{2\,000}{n}.$$

Dacă, de exemplu, $n=200\,000$, atunci

$$P(|k-150\,000|>2\,000)<0{,}01.$$

Din punct de vedere practic aceasta înseamnă, de exemplu, următoarele: dacă într'o ramură de producție oarecare cu un proces tehnologic stabilit 75% din produse, în medie, posedă o anumită proprietate (de exemplu sunt de calitatea I-a), atunci probabilitatea ca, din 200 000 produse, 148 000 până la 152 000 produse să posede această proprietate depășește 0,99 (adică acest eveniment este practic sigur).

In această problemă trebue să mai facem două observații:

1. Inegalitatea (8) dă o evaluare foarte grosolană pentru probabilitatea $P(|k-np|>\varepsilon n)$; în realitate, această probabilitate, mai ales când n este mare, este mult mai mică; în practică se folosesc din această cauză calcule mai precise, dar mai greu de demonstrat.

2. Evaluarea dată de inegalitatea (8) devine mult mai precisă dacă probabilitatea p este foarte mică, sau dacă ea este, dimpotrivă, foarte apropiată de unitate.

Astfel, dacă în exemplul de mai sus, probabilitatea ca produsul să posede o anumită proprietate este $p=0{,}95$, deci $1-p=0{,}05$, $p(1-p)<0{,}05$. Deaceea, punând $\varepsilon=0{,}005$, $n=200\,000$, obținem:

$$\frac{p(1-p)}{\varepsilon^2 n}<\frac{0{,}05\cdot 1\,000\,000}{25\cdot 200\,000}=0{,}01,$$

ca și mai sus. Acum însă εn nu este 2 000 ci numai 1 000; de unde (întrucât $np=190\,000$) concludem că practic este sigur că numărul de produse care posedă însușirea considerată, va fi cuprins între 189 000 și 191 000 dacă numărul total al produselor este de 200 000.

Astfel, pentru $p=0{,}95$, rezultă din inegalitatea (8) garanția practică că numărul de produse care are însușirea ce ne interesează este cuprins într'un interval cu o lungime de două ori mai scurtă, decât în cazul în care $p=0{,}75$, întrucât avem:

$$P(|k-190\,000|>1\,000)<0{,}01.$$

Scanned with OKEN Scanner

Problemă. Se știe că a patra parte din lucrătorii dintr'o ramură industrială oarecare au absolvit școala medie. Pentru verificare, s'au luat la întâmplare 200 000 lucrători. Să se afle: 1) valoarea cea mai probabilă a numărului de lucrători ce au absolvit școala medie dintre cei 200 000 aleși și 2) probabilitatea ca numărul adevărat efectiv al acestor lucrători să difere de numărul cel mai probabil, cu cel mult 1,6%.

In rezolvarea problemei, pornim dela faptul că probabilitatea de a avea școala medie absolvită este egală cu $1/4$ pentru fiecare dintre cei 200 000 de lucrători luați la întâmplare (aceasta este chiar sensul cuvântului „la întâmplare"). Astfel în problema noastră

$$n=200\,000,\quad p=\frac{1}{4},\quad k_0=np=50\,000;\quad p(1-p)=\frac{3}{16}.$$

Se caută probabilitatea ca $|(k-np)|<0{,}016\,np$ sau $|k-np|<800$, unde k este numărul lucrătorilor ce au absolvit școala medie.

Alegem ε astfel încât să avem $\varepsilon n=800$; rezultă de aici $\varepsilon=\frac{800}{n}=0{,}004$. Formula (8) dă:

$$P(|k-50\,000|>800)<\frac{3}{16\cdot 0{,}000\,016\cdot 200\,000}\approx 0{,}06,$$

de unde

$$P(|k-50\,000|<800)>0{,}94.$$

Răspuns. Valoarea cea mai probabilă căutată este 50 000, probabilitatea căutată este mai mare ca 0,94.

In realitate însă, probabilitatea căutată este mult mai apropiată de unitate.

PARTEA A DOUA

VARIABILE INTÂMPLĂTOARE

CAPITOLUL VII

VARIABILE INTÂMPLĂTOARE ȘI LEGI DE REPARTIȚIE

§ 18. Noțiunea de variabilă întâmplătoare

Am întâlnit până acum de multe ori mărimi a căror valoare numerică nu poate fi stabilită odată pentru totdeauna, ci se schimbă sub influența unor acțiuni întâmplătoare. Astfel, numărul băieților născuți la 100 de nou născuți nu va fi unul și același la fiecare dintre sute. Sau lungimea fibrelor de bumbac de o varietate determinată se schimbă nu numai dela un raion la altul, unde această varietate se cultivă, ci și la aceeași plantă și la aceeași capsulă. Mai citâm câteva exemple de astfel de mărimi.

1. Trăgând din aceeași armă în aceeași țintă și în condiții neschimbate, observăm totuși că proectilele cad în diferite locuri; acest fenomen se numește „dispersiunea" proectilelor. Distanța dintre locul de cădere a proectilelor și locul din care au pornit este o mărime care ia diferite valori numerice dela o tragere la alta, în funcție de circumstanțe *întâmplătoare*, care nu pot fi calculate în prealabil.

2. Viteza unei molecule de gaz nu rămâne neschimbată ci variază datorită ciocnirilor ei cu alte molecule. Intrucât fiecare moleculă poate sau să se ciocnească sau să nu se ciocnească cu o altă moleculă de gaz, schimbarea vitezei ei are un caracter *întâmplător*.

3. Numărul de meteorite care cad pe pământ în timp de un an, atingând suprafața lui, nu este constant, ci variază foarte mult, în funcție de o serie întreagă de circumstanțe, cu caracter *întâmplător*.

4. Greutatea boabelor de grâu, crescute pe un teren oarecare, nu are o valoare determinată, ci variază dela un bob la altul. Intrucât nu se poate evalua influența tuturor factorilor (calitatea

Scanned with OKEN Scanner

solului din terenul unde a crescut spicul cu bobul dat, condițiile de lumină ale bobului dat, regimul de apă și altele), care determină creșterea bobului, greutatea lui este o mărime care variază în funcție de *întâmplare*.

Cu toată diversitatea exemplelor examinate, toate, din punctul de vedere care ne interesează acum, se prezintă la fel; în fiecare dintre aceste exemple avem a face cu o mărime care caracterizează într'un fel sau altul rezultatul operației efectuate (de exemplu, numărarea meteoritelor, măsurarea lungimii fibrelor); fiecare dintre aceste mărimi poate să ia valori diferite, în operații diferite, oricât de omogene ar fi condițiile pe care le-am realiza, în funcție de deosebiri întâmplătoare în circumstanțele acestor operații, care scapă controlului nostru. In calculul probabilităților astfel de mărimi se numesc *variabile întâmplătoare*. Exemplele pe care le-am dat sunt suficiente pentru a ne convinge cât de important poate fi studiul variabilelor întâmplătoare în aplicarea calculului probabilităților la cele mai variate domenii ale științei și ale practicei.

A cunoaște o variabilă întâmplătoare dată nu înseamnă desigur a cunoaște valoarea ei numerică, căci dacă știm, de exemplu, că proectilul a căzut la 926 m distanță de punctul de plecare, prin aceasta chiar distanța respectivă a căpătat o valoare numerică determinată și a încetat de a mai fi o variabilă întâmplătoare. Ce trebue să știm despre o variabilă *întâmplătoare* pentru a avea toate datele posibile referitoare la ea, ca variabilă întâmplătoare? E clar că pentru aceasta trebue, în primul rând, să cunoaștem toate valorile numerice pe care această variabilă le poate lua. Astfel, dacă la o tragere dintr'un tun, în anumite condiții determinate, bătaia cea mai lungă observată a tunului a fost de 982 m, iar cea mai scurtă de 904 m, distanța dintre locul de cădere a proectilului, și locul de pornire poate să ia toate valorile cuprinse între aceste două limite. In exemplul 3 este clar că numărul de meteorite care au atins suprafața pământului în timp de un an poate fi orice număr întreg pozitiv, adică 0, 1, 2, 3, etc.

Totuși, numai cunoașterea mulțimii valorilor posibile ale variabilei întâmplătoare nu ne oferă încă acele date care să ne folosească ca material la evaluările practice necesare. Astfel, dacă în al doilea exemplu, vom studia gazul la două temperaturi diferite valorile numerice posibile ale vitezei moleculelor sunt aceleași, deci enumerarea acestor valori nu ne dă posibilitatea unei aprecieri comparative a acestor temperaturi.

Dar diferența dintre temperaturi indică o deosebire esențială în starea gazului, deosebire despre care simpla enumerare a

Scanned with OKEN Scanner

valorilor posibile ale vitezelor moleculelor nu ne poate da nicio imagine.

Dacă vrem să evaluăm temperatura unei mase date de gaz, iar nouă ni se dă numai lista valorilor posibile ale vitezei moleculelor sale, este firesc să ne întrebăm cât de frecvent se observă o viteză sau alta? Cu alte cuvinte vom căuta, în mod firesc, să aflăm probabilităţile diferitelor valori posibile ale variabilei *întâmplătoare* care ne interesează.

§ 19. Noţiunea de lege de repartiţie

Să considerăm deocamdată, un exemplu foarte simplu. Se face o tragere în panoul din fig 8. Nimerirea în regiunea *I* contează 3 puncte, în regiunea *II* — 2 puncte, şi în regiunea *III* — 1 punct [1].

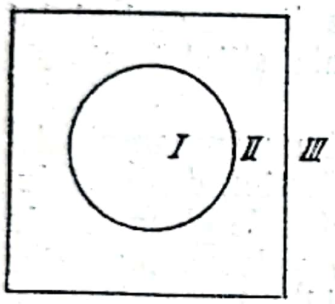

Fig. 8

Să luăm ca variabilă întâmplătoare numărul de puncte rezultate dintr'o tragere izolată; ca valori posibile figurează aici numerele 1, 2, 3; să însemnăm respectiv prin p_1, p_2, p_3, probabilităţile acestor trei valori astfel că p_3, de exemplu, reprezintă probabilitatea nimeririi în regiunea *I* din panou. In timp ce valorile posibile ale variabilei noastre întâmplătoare sunt aceleaşi pentru toţi trăgătorii, probabilităţile p_1, p_2 şi p_3 pot să difere mult dela un trăgător la altul şi se înţelege că prin această deosebire se defineşte şi deosebirea dintre calităţile tragerii. Astfel, pentru un trăgător foarte bun putem avea, de exemplu, $p_3=0{,}8$; $p_2=0{,}2$; $p_1=0$; pentru un trăgător mijlociu $p_3=0{,}3$; $p_2=0{,}5$; $p_1=0{,}2$; pentru un trăgător foarte slab $p_3=0{,}1$; $p_2=0{,}3$; $p_1=0{,}6$.

Dacă la o tragere oarecare se trag 12 focuri, ca valori posibile ale numărului de nimeriri în fiecare dintre regiunile *I*, *II* şi *III* vor servi toate numerele întregi dela 0 până la 12 inclusiv; dar acest fapt, prin el însuşi, nu ne dă încă posibilitatea de a aprecia calitatea tragerii; dimpotrivă, nu ne vom putea face o idee completă asupra calităţii acestei trageri decât dacă, pe lângă valorile posibile ale numărului de nimeriri în ţintă, ni se dau şi probabilităţile acestor valori, adică ni se dau numerele care arată cât de

[1] Cititorul poate obiecta că pentru nimerirea în regiunea *III* adică pentru ratări, nu se cuvine să se dea niciun punct. Totuşi, dacă 1 punct înseamnă dreptul de a trage odată, acela care a tras greşit a obţinut prin aceasta un punct.

Scanned with OKEN Scanner

frecvent se va întâlni în cele 12 focuri un număr sau altul de nimeriri într'o regiune sau alta.

Este clar că lucrurile se vor prezenta astfel în toate cazurile; cunoscând probabilităţile diferitelor valori posibile ale variabilei întâmplătoare, vom şti cât de des trebue să ne aşteptăm la apariţia valorilor ei mai mult sau mai puţin convenabile; iar aceasta, se înţelege, este suficient pentru a judeca gradul de eficacitate sau calitatea operaţiei de care este legată variabila întâmplătoare dată. Practica ne arată că cunoaşterea probabilităţilor tuturor valorilor posibile ale variabilei întâmplătoare studiate este într'adevăr suficientă pentru rezolvarea oricărei probleme legate de evaluarea acestei variabile întâmplătoare, ca fiind un indiciu calitativ al operaţiei respective.

Ajungem astfel la concluzia că, pentru caracterizarea completă a unei variabile întâmplătoare sau alteia, luată ca atare, este necesar şi suficient să se cunoască:

1. mulţimea valorilor posibile ale acestei variabile şi
2. probabilitatea fiecăreia dintre aceste valori.

Rezultă că este bine să se prezinte variabila întâmplătoare sub forma unei tabele cu două linii: linia de sus conţine într'o ordine oarecare toate valorile posibile ale variabilei întâmplătoare, iar linia de jos — probabilităţile lor, astfel încât sub fiecare dintre valorile posibile să se afle probabilitatea ei. Astfel, în exemplul de mai sus, numărul de puncte obţinut la un foc tras de către cel mai bun dintre trăgători poate fi reprezentat ca o variabilă întâmplătoare, prin tabela:

1	2	3
0	0,2	0,8

(1)

In general, variabila întâmplătoare care poate lua valorile x_1, x_2, ..., x_n, cu probabilităţile corespunzătoare p_1, p_2, ..., p_n se determină prin tabela:

x_1	x_2	...	x_n
p_1	p_2	...	p_n

A prezenta o astfel de tabelă, adică a da toate valorile posibile ale variabilei întâmplătoare împreună cu probabilităţile lor —

înseamnă, după cum se spune, a da *legea de repartiție* a acestei variabile întâmplătoare. Cunoașterea legii de repartiție a variabilei întâmplătoare date permite rezolvarea tuturor problemelor legate de ea.

Problemă. Numărul de puncte obținute la o singură tragere de către un trăgător are legea de repartiție (I); numărul de puncte pentru un alt trăgător are legea de repartiție:

1	2	3
0,2	0,5	0,3

(II)

Să se afle legea de repartiție a sumei punctelor realizate de amândoi trăgătorii.

Este evident că suma despre care este vorba este o variabilă întâmplătoare; noi trebue să-i formăm tabela. Pentru aceasta, trebue să examinăm toate rezultatele posibile ale tragerii concomitente a celor doi trăgători; vom pune aceste rezultate în tabela de mai jos, unde probabilitatea fiecărui rezultat se calculează conform regulii de înmulțire pentru evenimente independente și unde x reprezintă numărul de puncte realizate de primul trăgător, iar y — numărul de puncte realizate de cel de al doilea trăgător.

Nr. rezulatului	x	y	$x+y$	Probabilitatea rezultatului
1.	1	1	2	$0 \cdot 0{,}2=0$
2.	1	2	3	$0 \cdot 0{,}5=0$
3.	1	3	4	$0 \cdot 0{,}3=0$
4.	2	1	3	$0{,}2 \cdot 0{,}2=0{,}04$
5.	2	2	4	$0{,}2 \cdot 0{,}5=0{,}1$
6.	2	3	5	$0{,}2 \cdot 0{,}3=0{,}06$
7.	3	1	4	$0{,}8 \cdot 0{,}2=0{,}16$
8.	3	2	5	$0{,}8 \cdot 0{,}5=0{,}4$
9.	3	3	6	$0{,}8 \cdot 0{,}3=0{,}24$

Această tabelă ne arată că $x+y$, care este suma care ne interesează, poate lua valorile 3, 4, 5 și 6; valoarea 2 nu este posibilă în acest caz, întrucât probabilitatea ei este egală cu 0 [1]. Avem $x+y=3$ la rezultatele 2) și 4); deci, pentru ca suma $x+y$ să ia

[1] Putem desigur să considerăm numărul 2 ca valoare posibilă a variabilei $x+y$, având probabilitatea 0, tot așa cum pentru generalitate am făcut-o pentru valoarea 1 în tabela (I).

Scanned with OKEN Scanner

valoarea 3, trebue să se producă unul din rezultatele 2) sau 4); probabilitatea acestei valori, conform regulii de adunare, este egală cu suma probabilităţilor celor două rezultate, adică egală cu $0+0{,}04=0{,}04$. Pentru egalitatea

$$x+y=4$$

este necesar să se producă unul din rezultatele 3), 5) sau 7); probabilitatea acestei egalităţi este deci egală (tot conform regulii de adunare) cu $0+0{,}1+0{,}16=0{,}26$; în mod analog, găsim că probabilitatea valorii 5 pentru suma $x+y$ este

$$0{,}06+0{,}4=0{,}46,$$

iar probabilitatea valorii 6, care are loc numai în cazul rezultatului 9), este egală cu 0,24. Am obţinut astfel pentru variabila întâmplătoare $x+y$ următoarea tabelă a valorilor posibile:

3	4	5	6
0,04	0,26	0,46	0,24

(III)

Cu această tabelă (III) problema pusă este complect rezolvată.

Suma celor patru probabilităţi din tabela (III) este egală cu unu; aceasta este proprietatea pe care trebue s'o aibă orice lege de repartiţie, deoarece este vorba de suma probabilităţilor tuturor valorilor posibile ale variabilei întâmplătoare, adică de suma probabilităţilor unui sistem complect de evenimente oarecare. E util să folosim această proprietate a legilor de repartiţie ca un mijloc de a verifica calculele efectuate.

CAPITOLUL VIII

VALORI MEDII

§ 20. Definiţia valorii medii a unei variabile întâmplătoare

Cei doi trăgători despre care am vorbit mai înainte, trăgând simultan, pot realiza în funcţie de circumstanţe întâmplătoare, 3, 4, 5 sau 6 puncte; probabilităţile acestor patru rezultate posibile sunt date în tabela (III). Dacă ne punem întrebarea: „câte puncte realizează aceşti doi trăgători la un singur foc (dublu) tras?“ nu putem răspunde imediat deoarece pentru

Scanned with OKEN Scanner

diferitele focuri trase avem diferite rezultate. Dar pentru aprecierea calității tragerii celor doi trăgători, ne va interesa bineînțeles nu rezultatul unui foc izolat (acest rezultat poate fi întâmplător), ci rezultatul *mediu* al unei serii întregi de focuri. Câte puncte dă în medie un singur foc tras de perechea noastră de trăgători? Această întrebare are sens și poate primi un răspuns clar.

Să raționăm în felul următor. Dacă perechea noastră de trăgători trage 100 de focuri duble, atunci, după cum arată tabela (III),

aproximativ	4	din	aceste	focuri	vor	da	câte	3	puncte
„	26	„	„	„	„	„	„	4	„
„	46	„	„	„	„	„	„	5	„
„	24	„	„	„	„	„	„	6	„

Astfel, fiecare sută de focuri duble ale celor doi trăgători va da, în medie, un număr de puncte, care se exprimă prin suma

$$3\times4+4\times26+5\times46+6\times24=490.$$

Împărțind acest număr la 100, obținem că în medie la un foc tras revin 4,9 puncte; aceasta ne dă răspunsul la întrebarea pusă. Să observăm că, în loc de a împărți la 100 suma obținută (490) (cum am făcut aici), am putea, înainte de adunare să împărțim la 100 fiecare din termeni; atunci suma ne dă direct numărul mediu de puncte pentru un singur foc tras; cel mai simplu mod de a se efectua această împărțire este de a împărți la 100 al doilea factor al fiecărui termen al sumei: acești factori au fost doar obținuți prin înmulțirea probabilităților care figurează în tabela (III) cu 100 și pentru a le împărți cu 100 este suficient să revenim chiar la aceste probabilități. Pentru numărul mediu de puncte ce revine pentru un singur foc, obținem în felul acesta expresia:

$$3\times0{,}04+4\times0{,}26+5\times0{,}46+6\times0{,}24=4{,}9.$$

Suma ce se află în primul membru al acestei egalități este, după cum se vede imediat, construită cu datele din tabela (III), după o regulă foarte simplă: fiecare din valorile posibile indicate în prima linie a acestei tabele este înmulțită cu probabilitatea care se află în această tabelă dedesubtul ei, și apoi aceste produse sunt adunate.

Să facem același raționament într'o formă generală. Să presupunem că o variabilă întâmplătoare oarecare este dată în tabela

x_1	x_2	. . .	x_k
p_1	p_2	. . .	p_k

Scanned with OKEN Scanner

Reamintim: dacă probabilitatea valorii x_1 a variabilei x este p_1, aceasta înseamnă că, într'o serie de n operaţii, această valoare x_1 se va observa cam de n_1 ori unde $\frac{n_1}{n}=p_1$, de unde $n_1=np_1$; în mod analog, valoarea x_2 se va întâlni de aproximativ $n_2=np_2$ ori, etc. Astfel, şirul de n operaţii va conţine în medie

$$\begin{aligned} &n_1=np_1 \text{ operaţii, unde } x=x_1,\\ &n_2=np_2 \quad \text{„} \quad \text{„} \quad x=x_2,\\ &\cdots\cdots\cdots\cdots\\ &n_k=np_k \text{ operaţii, unde } x=x_k. \end{aligned}$$

Suma valorilor variabilei x în toate cele n operaţii efectuate va fi deci aproximativ egală cu

$$x_1n_1+x_2n_2+\ldots+x_kn_k=n(x_1p_1+x_2p_2+\ldots+x_kp_k).$$

Deaceea valoarea medie a variabilei întâmplătoare $\bar{x}$, care corespunde unei operaţii separate şi se obţine din suma scrisă împărţită cu numărul n al operaţiilor din seria dată, va fi

$$\bar{x}=x_1p_1+x_2p_2+\ldots+x_kp_k.$$

Am obţinut în felul acesta următoarea regulă importantă: *Pentru a obţine valoarea medie a unei variabile întâmplătoare trebue să înmulţim fiecare dintre valorile ei posibile cu probabilităţile corespunzătoare lor şi să adunăm între ele toate produsele obţinute.*

La ce ne poate folosi cunoaşterea valorii medii a unei variabile întâmplătoare? Pentru a răspunde mai precis la această întrebare, să mai examinăm câteva exemple.

Exemplul 1. Să revenim încă odată la cei doi trăgători; punctele pe care ei le realizează sunt variabile întâmplătoare ale căror legi de repartiţie sunt date în tabela (I) pentru primul trăgător şi în tabela (II), pentru cel de al doilea (pag. 57 şi 58). O cercetare atentă a acestor două tabele ne arată imediat că primul trăgător trage mai bine decât al doilea; în adevăr probabilitatea celui mai bun rezultat (3 puncte) este mult mai mare la primul decât la al doilea trăgător, pe când probabilităţile rezultatelor mai slabe sunt mai mari la cel de al doilea decât la primul. Totuşi, o astfel de comparaţie nu ne satisface, căci ea are un caracter pur calitativ, nu ne dă acea măsură, acel număr, care prin mărimea lui să indice în mod direct calitatea tragerii unuia sau celuilalt trăgător după cum temperatura, de exemplu, indică direct gradul de încălzire a unui corp. Fără să avem o astfel de măsură de apreciere, ne putem găsi într'o situaţie în care examinarea nemijlocită nu ne va

Scanned with OKEN Scanner

da nici un răspuns sau în care acest răspuns s'ar dovedi discutabil. Astfel, dacă în locul tabelelor (I) și (II) am avea tabelele:

1	2	3
0,4	0,1	0,5

(I') și

pentru primul trăgător

1	2	3
0,1	0,6	0,3

(II')

pentru al doilea trăgător

ar fi greu să hotărîm dintr'o singură privire aruncată asupra acestor tabele, care dintre cei doi trăgători trage mai bine; desigur un rezultat mai bun (3 puncte) este mai probabil pentru primul decât pentru cel de al doilea; dar totodată și cel mai prost rezultat (1 punct) este deasemenea mai probabil pentru primul decât pentru al doilea; dimpotrivă, rezultatul mijlociu (cu două puncte) este mai probabil la al doilea decât la primul.

Să formăm acum, după regula de mai sus, valoarea medie a numărului de puncte pentru fiecare dintre cei doi trăgători.

1. pentru primul trăgător:

$$1\times0,4+2\times0,1+3\times0,5=2,1;$$

2. pentru al doilea trăgător:

$$1\times0,1+2\times0,6+3\times0,3=2,2.$$

După cum se vede, al doilea trăgător are în medie un număr de puncte ceva mai mare decât primul; aceasta înseamnă practic că la o tragere repetată al doilea trăgător va da, în general, un rezultat ceva mai bun decât primul trăgător. Putem afirma acum cu certitudine că al doilea trăgător trage mai bine la țintă; valoarea medie a numărului de puncte realizate ne-a dat o măsură comodă de comparare a măestriei celor doi trăgători, deoarece este o metodă ușoară și care nu lasă nici un dubiu.

Exemplul 2. La montarea unui aparat de precizie pentru o ajustare mai precisă a unei anumite piese se pot cere, în funcție de reușită, 1, 2, 3, 4 sau 5 probe. Astfel, numărul x de probe necesare pentru o montare satisfăcătoare este o variabilă întâmplătoare care poate lua valorile 1, 2, 3, 4, 5; să presupunem, de exemplu, că probabilitățile acestor valori sunt date în tabela

1	2	3	4	5
0,07	0,16	0,46	0,21	0,1

Scanned with OKEN Scanner

Putem fi puşi în faţa problemei de a aproviziona pe lucrătorul care execută montarea cu un număr de piese de un anumit fel, necesare pentru 20 aparate [1]. Pentru a avea posibilitatea de a ne orienta în aprecierea acestui număr, nu vom putea folosi direct tabela dată, care ne spune doar cum se petrec lucrurile dela caz la caz. Dacă însă vom găsi valoarea medie $\bar{x}$ a numărului x de probe necesare pentru un aparat şi vom înmulţi această valoare medie cu 20, vom obţine, după cum se vede clar, o valoare a numărului dat cu care ne putem orienta. Obţinem

$$\bar{x}=1\times 0{,}07+2\times 0{,}16+3\times 0{,}46+4\times 0{,}21+5\times 0{,}1=3{,}11;$$

$$20\bar{x}=3{,}11\times 20=62{,}2\approx 62.$$

Pentru ca montorul să aibă o mică rezervă pentru cazul când consumul efectiv de piese ar depăşi pe cel aşteptat, ar fi util în practică să-i dăm 65—70 piese.

In exemplele examinate, am avut a face cu situaţii în care se cerea în practică o anumită evaluare a unei variabile întâmplătoare care să ne orienteze; examinând doar tabela nu putem face o astfel de evaluare — căci tabela ne arată numai că variabila întâmplătoare poate lua anumite valori cu anumite probabilităţi. Dar *valoarea medie* a variabilei întâmplătoare, calculată după această tabelă, ne poate permite o astfel de evaluare, deoarece ea reprezintă valoarea pe care variabila noastră o va lua în medie la un număr suficient de mare de operaţii. Vedem că, din punct de vedere practic, valoarea medie caracterizează foarte bine variabila întâmplătoare, atunci când este vorba de o operaţie de masă, sau repetată de mai multe ori.

Problema 1. Se face o serie de experienţe cu una şi aceeaşi probabilitate p de producere a unui eveniment oarecare A, rezultatele diferitelor experienţe fiind independente între ele. Să se afle valoarea medie a numărului de produceri ale evenimentului A într'o serie de n probe.

Numărul de produceri ale evenimentului A într'o serie de n probe este o variabilă întâmplătoare ce poate lua valorile 0,1, 2, ..., n, probabilitatea valorii k fiind egală, după cum ştim, cu

$$P_n(k)=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}.$$

Deci, valoarea medie căutată este

$$\sum_{k=0}^{n} k\,P_n(k).$$

[1] Vom presupune, în acest caz, că o piesă rebutată la montarea unui aparat, nu mai este folosită la montarea altor aparate.

Scanned with OKEN Scanner

Am calculat această sumă cu ocazia demonstrației teoremei lui Bernoulli (pag. 50) și am văzut că este egală cu *np*. Ne-am convins la timpul său că numărul *cel mai probabil* de produceri ale evenimentului *A* în *n* probe, *n* fiind un număr mare, este apropiat de *np*; vedem acum că numărul *mediu* de produceri ale evenimentului *A*, oricare ar fi *n*, este exact egal cu *np*. Astfel, în cazul dat, valoarea cea mai probabilă a variabilei întâmplătoare coincide cu valoarea ei medie; nu trebue totuși să credem că această coincidență are loc pentru orice variabilă întâmplătoare; în general, cea mai probabilă valoare a variabilei întâmplătoare poate fi situată foarte departe de valoarea ei medie. Astfel, de exemplu, pentru variabila întâmplătoare cu legea de repartiție

0	5	10
0,7	0,1	0,2

valoarea cea mai probabilă este 0, iar valoarea medie este 2,5.

Problema 2. Se trage un număr de focuri cu probabilitatea de nimerire 0,8; tragerea se face până la prima nimerire la țintă, dar într'un număr de cel mult 4 focuri; să se determine valoarea medie a numărului de focuri trase.

Numărul de focuri ce va trebui să se tragă poate fi egal, conform condițiilor tragerii, cu 1, 2, 3 sau 4. Trebue să calculăm probabilitățile acestor 4 valori. Pentru ca să fie nevoie de un singur foc, trebue ca chiar primul foc să nimerească ținta; probabilitatea pentru aceasta este

$$p_1 = 0{,}8.$$

Pentru ca să fie nevoie de două focuri, trebue ca primul foc să se dea greș, iar al doilea să nimerească ținta.

Probabilitatea aceasta, conform regulii de înmulțire pentru evenimente independente, este:

$$p_2 = (1-0{,}8) \times 0{,}8 = 0{,}16.$$

Pentru ca să fie nevoie de trei focuri, trebuie să se rateze ținta la primele două focuri, iar la al treilea — să se nimerească. Deaceea

$$p_3 = (1-0{,}8)^2 \times 0{,}8 = 0{,}032.$$

Insfârșit, sunt necesare 4 focuri, dacă primele trei vor fi ratate (indiferent de ceeace se va întâmpla la al patrulea foc); deaceea

$$p_4 = (1-0{,}8)^3 = 0{,}008.$$

Scanned with OKEN Scanner

Astfel, numărul tragerilor efectuate, ca variabilă întâmplătoare, este determinat de legea de repartiție

1	2	3	4
0,8	0,16	0,032	0,008

Valoarea medie a acestui număr este deci

$$1\times0{,}8+2\times0{,}16+3\times0{,}032+4\times0{,}008=1{,}248.$$

Dacă de exemplu trebue să se execute dintr'un tun 100 asemenea trageri, ne putem aștepta că se va cheltui, în acest caz, aproximativ $1{,}248\times100\approx125$ obuze.

Problema 3. Un teren oarecare are forma unui pătrat, având latura egală, după datele măsurătorilor aerofotometrice, de 350 m. Calitatea aerofotografierii se determină prin aceea că eroarea de

0 m are probabilitatea de 0,42
±10 m „ „ „ 0,16 [1]
±20 m „ „ „ 0,08
±30 m „ „ „ 0,05

Să se afle valoarea medie a suprafeței terenului.

In funcție de condițiile întâmplătoare ale măsurării aerofotometrice, latura pătratului este o variabilă întâmplătoare a cărei lege de repartiție este prezentată în tabela:

320	330	340	350	360	370	380
0,05	0,08	0,16	0,42	0,16	0,08	0,05

(1)

De aici am putea să aflăm dintr'odată valoarea medie a acestei mărimi; în acest caz, nu este nici măcar nevoie să folosim regula noastră de calcul; în adevăr, întrucât erorile egale într'un sens și altul sunt la fel de probabile, din simetria lor este clar că valoarea medie a laturii pătratului este egală cu mărimea

[1] Prin aceasta se înțelege că eroarea de $+10$ m și eroarea de -10 m au fiecare probabilitatea de 0,16; la fel trebue să înțelegem și probabilitățile date mai jos.

observată, adică 350 m; mai amănunțit: expresia valorii medii va conține termenii:

$$(340+360)\cdot 0{,}16=(350+10)+(350-10)\cdot 0{,}16=2\cdot 350\cdot 0{,}16,$$
$$(330+370)\cdot 0{,}08=2\cdot 350\cdot 0{,}08,$$
$$(320+380)\cdot 0{,}05=2\cdot 350\cdot 0{,}05;$$

și este deci egală cu $350\cdot(0{,}42+2\cdot 0{,}16+2\cdot 0{,}08+2\cdot 0{,}05)=350$.

S'ar putea crede că, pe baza aceloraşi considerente de simetrie, valoarea medie a suprafeței pătratului ar trebui să fie $350^2=122\,500$ m²; așa ar fi, dacă valoarea medie a pătratului unei variabile întâmplătoare ar fi egală cu pătratul valorii ei medii; nu este însă așa; în exemplul nostru, suprafața pătratului poate avea valorile

$$320^2,\ 330^2,\ 340^2,\ 350^2,\ 360^2,\ 370^2,\ 380^2;$$

care din aceste valori are loc în realitate, depinde de care caz din cele șapte prezentate în tabela (I) are loc, așa că probabilitățile acestor șapte valori sunt cele din tabela (I); pe scurt, legea de repartiție a suprafeței pătratului este dată de tabela:

320^2	330^2	340^2	350^2	360^2	370^2	380^2
0,05	0,08	0,16	0,42	0,16	0,08	0,05

deci valoarea ei medie este

$$320^2\cdot 0{,}05+330^2\cdot 0{,}08+340^2\cdot 0{,}16+350^2\cdot 0{,}42,+360^2\cdot 0{,}16+$$
$$+370^2\cdot 0{,}08+380^2\cdot 0{,}05.$$

Este folositor și în acest caz ca, pentru simplificarea calculelor, să folosim simetria existentă; este util să vedem cum se face acest lucru, întrucât asemenea posibilități de simplificare apar destul de des. Putem scrie expresia de mai sus sub forma

$$350^2\cdot 0{,}42+(340^2+360^2)\cdot 0{,}16+(330^2+370^2)\cdot 0{,}08+$$
$$+(320^2+380^2)\cdot 0{,}05=$$
$$=350^2\cdot 0{,}14+[(350-10)^2+(350+10)^2]\cdot 0{,}16+$$
$$+[(350-20)^2+(350+20)^2]\cdot 0{,}08+$$
$$+[(350-30)^2+(350+30)^2]\cdot 0{,}05=$$
$$=350^2[0{,}42+2\cdot 0{,}16+2\cdot 0{,}08+2\cdot 0{,}05]+$$
$$+2\cdot 10^2\cdot 0{,}16+2\cdot 20^2\cdot 0{,}08+2\cdot 30^2\cdot 0{,}05=$$
$$=350^2+2\cdot(16+32+45)=122\,686.$$

Cu această metodă toate calculele pot fi făcute mintal.

Vedem că valoarea medie a suprafeței pătratului este ceva mai mare (este adevărat că în practică diferența în acest caz este aproape imperceptibilă), decât pătratul valorii medii a laturii

Scanned with OKEN Scanner

(adică decât $350^2 = 122\,500$); este ușor de demonstrat că la baza acestui fapt stă următoarea regulă generală: valoarea medie a pătratului oricărei variabile întâmplătoare este întotdeauna mai mare decât pătratul valorii ei medii. In adevăr, fie o variabilă întâmplătoare x cu o lege de repartiție absolut arbitrară.

x_1	x_2	...	x_k
p_1	p_2	...	p_k

mărimea x^2 va avea atunci legea de repartiție

x_1^2	x_2^2	...	x_k^2
p_1	p_2	...	p_k

valorile medii ale acestor două mărimi sunt respectiv egale cu

$$\bar{x} = x_1 p_1 + x_2 p_2 + \ldots + x_k p_k$$

și

$$\bar{x^2} = x_1^2 p_1 + x_2^2 p_2 + \ldots + x_k^2 p_k.$$

Avem însă

$$\bar{x^2} - (\bar{x})^2 = \bar{x^2} - 2(\bar{x})^2 + (\bar{x})^2;$$

intrucât $p_1 + p_2 + \ldots + p_k = 1$, cei trei termeni ai membrului doi pot fi scriși respectiv sub forma

$$\bar{x^2} = \sum_{i=1}^{k} x_i^2 p_i,$$

$$2(\bar{x})^2 = 2(\bar{x})(\bar{x}) = 2(\bar{x}) \sum_{i=1}^{k} x_i p_i = \sum_{i=1}^{k} 2\bar{x}\, x_i p_i,$$

$$(\bar{x})^2 = (\bar{x})^2 \sum_{i=1}^{k} p_i = \sum_{i=1}^{k} (\bar{x})^2 p_i;$$

deaceea

$$\bar{x^2} - (\bar{x})^2 = \sum_{i=1}^{k} \{x_i^2 - 2\bar{x}\, x_i + (\bar{x})^2\} p_i = \sum_{i=1}^{k} (x_i - \bar{x})^2 p_i;$$

deoarece toți termenii sumei din membrul doi nu sunt negativi

$$\bar{x^2} - (\bar{x})^2 > 0,$$

ceeace trebuia demonstrat.

Scanned with OKEN Scanner

CAPITOLUL IX

VALORILE MEDII ALE SUMEI ȘI ALE PRODUSULUI

§ 21. Teorema valorii medii a sumei

Deseori se ivește necesitatea de a calcula valoarea medie a sumei a două (și uneori a mai multe) variabile întâmplătoare, cunoscând valorile medii ale fiecăreia. Să considerăm de exemplu, două întreprinderi care fabrică aceleași produse, fiind cunoscut că, în medie, prima întreprindere fabrică zilnic 120 de produse, iar a doua 180; putem noi, cu ajutorul acestor date, să stabilim valoarea medie a numărului de produse la care trebue să ne așteptăm zilnic din partea celor două întreprinderi luate împreună? Sau poate aceste date sunt insuficiente, și pe lângă valorile medii, mai avem nevoie să cunoaștem încă ceva despre cele două variabile întâmplătoare considerate (de exemplu, să cunoaștem în întregime legile lor de repartiție)?

Este foarte important faptul că, pentru a calcula valoarea medie a sumei, este întotdeauna suficient să cunoaștem valorile medii ale termenilor ei și că valoarea medie a sumei se exprimă prin valorile medii ale termenilor sumei sub forma cea mai simplă posibilă: *valoarea medie a sumei este întotdeauna egală cu suma valorilor medii ale termenilor ei.*

Astfel, dacă x și y sunt două variabile întâmplătoare absolut arbitrare, avem

$$\overline{x+y}=\overline{x}+\overline{y}.$$

In exemplul de mai sus, x este numărul de produse al primei întreprinderi, y, numărul de produse al întreprinderii a doua: $\overline{x}=120$, $\overline{y}=180$, deci

$$\overline{x+y}=\overline{x}+\overline{y}=300.$$

Pentru a demonstra într'o formă generală regula enunțată, să presupunem că variabilele x și y sunt supuse respectiv legilor de repartiție:

x_1	x_2	...	x_k
p_1	p_2	...	p_k

(I) și

y_1	y_2	...	y_l
q_1	q_2	...	q_l

(II)

Atunci valorile posibile ale variabilei $x+y$ vor fi toate sumele posibile de forma x_i+y_j unde $1 \leq i \leq k$ și $1 \leq j \leq l$. Probabilitatea

Scanned with OKEN Scanner

valorii x_i+y_j, pe care o vom însemna prin p_{ij} ne este necunoscută; aceasta este probabilitatea evenimentului dublu $x=x_i$ și $y=y_j$, adică probabilitatea ca variabila x să ia valoarea x_i, iar variabila y valoarea y_j. Dacă am presupune că aceste două evenimente sunt independente, am obține evident, conform regulii de înmulțire,

$$p_{ij}=p_i q_j. \qquad (1)$$

Noi nu vom presupune însă independența acestor evenimente.

Astfel, egalitatea (1) nu va fi valabilă și va trebui să ținem seama că simpla cunoaștere a tabelelor (I) și (II) nu permite să se tragă concluzii referitoare la mărimile p_{ij}.

Conform regulii generale, valoarea medie a variabilei $x+y$ este egală cu suma produselor tuturor valorilor posibile ale acestei variabile cu probabilitățile respective, adică

$$\overline{x+y}=\sum_{i=1}^{k}\sum_{j=1}^{l}(x_i+y_j)\,p_{ij}=\sum_{i=1}^{k}x_i\left(\sum_{j=1}^{l}p_{ij}\right)+\sum_{j=1}^{l}y_j\left(\sum_{i=1}^{k}p_{ij}\right). \qquad (2)$$

Să examinăm acum mai atent suma $\sum_{j=1}^{l}p_{ij}$; aceasta este suma probabilităților tuturor evenimentelor posibile de forma $(x=x_i,\ y=y_j)$ unde indicele i este unul și același pentru toți termenii sumei, iar indicele j variază dela un termen la altul, parcurgând tot șirul valorilor posibile (dela 1 până la l inclusiv); deoarece evenimentele $y=y_j$ pentru diferiți j sunt, evident, incompatibile, conform regulii de adunare, suma $\sum_{j=1}^{l}p_{ij}$ este probabilitatea producerii *unui eveniment oarecare din l evenimente* $(x=x_i,\ y=y_j)$ $(j=1, 2, \ldots, i)$.

Dar a spune „s'a produs unul oarecare din evenimentele $x=x_i$, $y=y_j\ (1\leq j\leq l)$" — este echivalent cu a spune numai „s'a produs evenimentul $x=x_i$"; într'adevăr: 1) dacă s'a produs unul din evenimentele $(x=x_i,\ y=y_j)$ (oricare ar fi j), este evident că s'a produs și evenimentul $x=x_i$; 2) reciproc, dacă s'a produs evenimentul $x=x_i$, atunci, întrucât y ia neapărat una din valorile sale posibile $y_1, y_2, \ldots, y_l$; trebue să se producă și unul oarecare din evenimentele $(x=x_i,\ y=y_j)\ (1\leq j\leq l)$. Astfel $\sum_{j=1}^{l}p_{ij}$ fiind probabilitatea producerii unuia oarecare dintre evenimentele $(x=x_i,\ y=y_j)\ (1\leq j\leq l)$ este egală cu probabilitățile evenimentului $x=x_i$, adică

$$\sum_{j=1}^{l}p_{ij}=p_i.$$

Scanned with OKEN Scanner

In mod analog ne convingem că

$$\sum_{i=1}^{k} p_{ij} = q_j ;$$

introducând aceste expresii în egalitatea (2), obținem

$$\overline{x+y} = \sum_{i=1}^{k} x_i p_i + \sum_{j=1}^{l} y_j q_j = \bar{x} + \bar{y},$$

ceeace trebuia demonstrat.

Teorema demonstrată de noi pentru cazul a doi termeni ai sumei se extinde în mod automat la cazul a trei sau mai mulți termeni; în adevăr, în virtutea celor demonstrate aici, putem scrie

$$\overline{x+y+z} = \overline{x+y}+\bar{z} = \bar{x}+\bar{y}+\bar{z}$$

etc.

Exemplu. Se trag n focuri cu probabilitățile de nimerire p_1, p_2, ..., p_n; se cere să aflăm valoarea medie a numărului total de nimeriri.

In cazul unui foc izolat, numărul de nimeriri este o variabilă întâmplătoare care poate lua numai valoriie 1 (nimerire) și 0 (ratare); probabilitățile acestor valori pentru primul foc sunt egale respectiv cu p_1 și $1-p_1$, din care cauză valoarea medie a numărului de nimeriri la primul foc este

$$1 \cdot p_1 + 0 \cdot (1-p_1) = p_1 .$$

Pentru al doilea foc, ea este egală cu p_2 și așa mai departe. Numărul total de nimeriri este suma numerelor de nimeriri ale focurilor separate; valoarea sa medie este deci egală, conform regulii de adunare a valorilor medii stabilită mai sus, cu suma valorilor medii ale numerelor de nimeriri corespunzătoare focurilor separate, adică este egală cu

$$p_1 + p_2 + \ldots + p_n ,$$

și astfel problema este rezolvată.

In particular, dacă posibilitatea nimeririi este aceeași pentru toate focurile trase ($p_1 = p_2 = \ldots = p_n = p$), valoarea medie a numărului total de nimeriri este np. Am obținut mai înainte acest rezultat la pag. 63 (problema 1). Este interesant să se compare calculele greoaie, pe care a trebuit să le facem acolo, cu raționamentul simplu și care nu cere nici un fel de calcul, care ne-a dus aici la același rezultat. Dar noi n'am simplificat numai lucrurile ci am făcut și o generalizare. In cazul rezultatului anterior, am presupus rezultatele diferitelor focuri ca fiind evenimente reciproc independente și, în adevăr, metoda folosită acolo era valabilă numai în această ipoteză; acum însă ne putem lipsi de această ipoteză întrucât regula de adunare a valorilor medii pe

Scanned with OKEN Scanner

care ne bazăm în noua demonstrație este valabilă pentru orice variabile întâmplătoare fără nicio restricție. Astfel, oricare ar fi dependența reciprocă dintre diferitele focuri, este suficient ca probabilitatea de nimerire p să fie una și aceeași pentru toate focurile trase, pentru ca valoarea medie a numărului total de nimeriri să fie întotdeauna egală cu np.

§ 22. Teorema valorii medii a produsului

Aceeași problemă pe care am rezolvat-o pentru sumele de variabile întâmplătoare se prezintă deseori și pentru produsele lor. Să presupunem că variabilele întâmplătoare x și y sunt supuse din nou respectiv legilor de repartiție indicate în tabelele (I) și (II). Produsul xy este atunci o variabilă întâmplătoare ale cărei valori posibile sunt produsele de formă $x_i y_j$ $(1 \leq i \leq k,\ 1 \leq j \leq l)$ probabilitatea valorii $x_i y_j$ este p_{ij}. Problema constă în găsirea unei reguli care să ne permită să exprimăm în toate cazurile valoarea medie $\overline{xy}$ a mărimii xy cu ajutorul valorilor medii ale factorilor. Însă rezolvarea acestei probleme este, în general, imposibilă. Cunoașterea valorilor medii $\overline{x}$ și $\overline{y}$ nu determină, în general, valoarea $\overline{xy}$ în mod unic (adică pentru unele și aceleași $\overline{x}$ și $\overline{y}$ sunt posibile diferite valori pentru mărimea $\overline{xy}$); din această cauză nu poate exista nici o formulă care să exprime $\overline{xy}$ prin $\overline{x}$ și $\overline{y}$.

Există însă un singur caz important când o astfel de expresie este posibilă și atunci relația obținută are o formă foarte simplă.

Convenim să numim variabilele întâmplătoare x și y *mutual independente* dacă evenimentele $x=x_i$ și $y=y_j$ sunt reciproc independente, pentru orice i și j, adică dacă condiția ca una din cele două variabile întâmplătoare să capete o valoare determinată nu influențează legea de repartiție a celei de a doua variabile întâmplătoare. Dacă mărimile x și y sunt mutual independente în sensul definit mai sus, atunci

$$p_{ij}=p_i q_j \ (i=1,\ 2,\ \ldots,\ k;\ j=1,\ 2,\ \ldots,\ l)$$

conform regulii de înmulțire pentru evenimente independente; deaceea

$$\overline{xy}=\sum_{i=1}^{k}\sum_{j=1}^{l} x_i y_j p_{ij}=\sum_{i=1}^{k}\sum_{j=1}^{l} x_i y_j p_i q_j=\sum_{i=1}^{k} x_i p_i \sum_{j=1}^{l} y_i q_j=\overline{x}\cdot\overline{y};$$

pentru variabile întâmplătoare mutual independente, valoarea medie a produsului este egală cu produsul valorilor medii ale factorilor.

Scanned with OKEN Scanner

Ca și în cazul adunării, această regulă pe care am dedus-o pentru produsul a două variabile întâmplătoare se extinde imediat pentru produsul unui număr oarecare de factori; este necesar, în acest caz, ca factorii să fie mutual independenți, adică alegerea unor valori determinate pentru unele din aceste variabile să nu influențeze legile de repartiție ale celorlalte variabile.

Exemplul 1. Să presupunem că se cere să măsurăm aria unui teren dreptunghiular prin aerofotografiere și că, din măsurarea laturilor acestui dreptunghi, a rezultat că ele au respectiv 72 m și 50 m. Legea de repartiție a erorilor măsurătorii este necunoscută, dar se știe că erorile de aceeași valoare în ambele sensuri sunt egal probabile; se înțelege atunci că, din considerente de simetrie (și se poate demonstra ușor v. problema 3, pag. 65), valorile medii ale laturilor dreptunghiului coincid cu rezultatele obținute la măsurare. Dacă aceste două rezultate ale măsurătorii pot fi considerate ca variabile întâmplătoare mutual independente, valoarea medie a suprafeței, conform regulii de înmulțire dedusă mai sus, va fi egală cu produsul valorilor medii ale laturilor, adică $72 \times 50 = 3\,600$ m². Uneori însă va trebui să presupunem că măsurătorile laturilor sunt mutual dependente. Aceasta va avea loc, de exemplu, în cazul când aceste două măsurători se fac cu aceleași instrumente, insuficient puse la punct. Dacă măsurarea lungimii dă un rezultat care depășește în mod simțitor adevărata lungime, avem în mod natural, dreptul să presupunem că instrumentele de măsurare au în general tendința de a arăta mărimi prea mari, din care cauză crește și probabilitatea ca măsurarea lățimii să dea mărimi exagerate, așa că aceste două mărimi nu pot fi socotite mutual independente. In asemenea cazuri, valoarea medie a suprafeței nu poate fi considerată egală cu produsul valorilor medii ale laturilor dreptunghiului și pentru determinarea acestei suprafețe se cer date suplimentare.

Exemplul 2. Printr'un conductor, a cărui rezistență depinde de împrejurări întâmplătoare, trece un curent electric a cărui intensitate depinde deasemenea de întâmplare. Se știe că valoarea medie a rezistenței conductorului este de 25 ohmi, iar intensitatea medie a curentului este de 6 amperi. Se cere să se calculeze valoarea medie a forței electromotrice E a curentului ce trece prin conductor.

Conform legii lui Ohm

$$E = RI,$$

unde R este rezistența conductorului, iar I intensitatea curentului.

Scanned with OKEN Scanner

Intrucât, conform celor presupuse de noi,

$$\bar{R}=25,\ \bar{I}=6,$$

presupunând mărimile R și I mutual independente, obținem

$$E=RI=25\cdot 6=150 \text{ volți.}$$

CAPITOLUL X

DISPERSIE ȘI ABATERI MEDII

§ 23. Insuficiența valorii medii pentru caracterizarea unei variabile întâmplătoare

Am văzut în repetate rânduri că valoarea medie a unei variabile întâmplătoare ne dă o primă imagine aproximativă, de orientare, a acestei variabile și că, în multe cazuri, această imagine este suficientă pentru scopul practic pe care-l urmărim. Astfel, pentru a compara gradul de măestrie a doi trăgători în întrecere, ne-a fost suficient să cunoaștem valorile medii ale numărului de puncte realizate de ei; pentru a compara eficacitatea a două sisteme diferite de minare a unui sector dat, este suficient să se cunoască valorile medii ale pagubelor pe care aceste două sisteme le pot pricinui dușmanului, etc. In toate aceste cazuri, avem un mare avantaj căci putem caracteriza variabila întâmplătoare printr'un număr unic — valoarea ei medie — în loc să o exprimăm prin complicata ei lege de repartiție. Lucrurile se petrec ca și când am avea a face nu cu o variabilă întâmplătoare, ci cu o variabilă cu o valoare bine determinată.

Mai des însă se întâlnește cazul în care trăsăturile variabilei întâmplătoare, cele mai importante pentru practică, nu sunt în nicio măsură determinate de valoarea ei medie, ci necesită o cunoaștere mai amănunțită a legii ei de repartiție. Un caz tipic de acest fel se prezintă la examinarea lungimii de sbor a proectilului. Să presupunem că x este distanța între locul de cădere a proectilului și locul din care a pornit. In cazul celei mai bune reglări a tunului, valoarea medie $\bar{x}$ a acestei distanțe este egală cu distanța dela tun până la țintă; să presupunem că $\bar{x}$ ne este cunoscut și că s'a realizat această reglare optimă a tunului. Cum vor cădea proectilele? Multe dintre ele vor atinge ținta? Sau vor cădea cel puțin aproape de ea? La toate aceste probleme care se pun, cunoscând numai pe $\bar{x}$, nu putem da nici un răspuns; știm

numai că vor cădea proectile mai departe decât ţinta sau mai aproape şi că şansele de cădere dincoace şi dincolo de ţintă sunt aproximativ aceleaşi, pentrucă valoarea medie a lungimii traiectoriei coincide cu distanţa până la ţintă. Nu cunoaştem însă principalul: vor cădea proectilele în majoritatea lor aproape de centrul ţintei, ca să ne putem aştepta la un număr mare de lovituri în plin, sau cea mai mare parte a proectilelor vor fi împrăştiate de ambele părţi ale ţintei şi la mare distanţă de ea, aşa încât numărul loviturilor în plin să fie nesatisfăcător. Este clar că ambele ipoteze se pot realiza pentru orice valoare aleasă pentru x.

Două tunuri având una şi aceeaşi valoare medie x, se pot comporta sub acest raport în moduri cu totul diferite; unul dintre ele poate da „o dispersie" mult mai mare de proectile decât celălalt. Aceasta înseamnă că, la tragerea din primul tun, proectilele vor cădea, în medie, mult mai departe de ţintă, de ambele părţi ale acesteia, decât la tragerea din al doilea tun, care se dovedeşte astfel mai bun, cu toate că valoarea x este aceeaşi pentru ambele tunuri.

Iată un alt exemplu. Să presupunem că se face proba fertilităţii a două specii de grâu. In funcţie de circumstanţe întâmplătoare (cantitatea precipitaţiilor, repartiţia îngrăşămintelor, radiaţia solară, etc.), recolta la un metru pătrat este supusă unor oscilaţii simţitoare şi reprezintă o variabilă întâmplătoare. Să presupunem că, în condiţii identice, recolta medie este aceeaşi pentru fiecare specie de grâu şi anume 240 g la un metru pătrat. Putem să ne facem o idee despre calitatea speciei de grâu încercat numai după valoarea recoltei medii? Desigur că nu, pentru că cel mai mare interes economic îl prezintă specia a cărei fertilitate este mai puţin supusă influenţelor întâmplătoare ale factorilor meteorologici şi ale altor factori, cu alte cuvinte, acea specie care are „dispersia" fertilităţii mai mică. Vedem astfel că, la încercarea fertilităţii unei specii de grâu, o importanţă tot atât de mare ca şi fertilitatea medie o are amplitudinea oscilaţiilor posibile.

§ 24. Diferite metode de măsurare a dispersiei unei variabile întâmplătoare

Exemplele citate mai sus precum şi numeroase alte exemple analoage ne arată în mod convingător că, în multe cazuri, pentru descrierea particularităţilor celor mai interesante, din punct de vedere practic, ale variabilelor întâmplătoare, cunoaşterea valorilor lor medii este cu totul insuficientă. Aceste trăsături interesante din punct de vedere practic rămân, în cazul când sunt

Scanned with OKEN Scanner

prezentate numai prin valorile lor medii, absolut necunoscute și pentru lămurirea lor este necesar să avem în fața noastră întreagă tabelă a repartiției unei astfel de variabile, ceeace practic este întotdeauna complicat și incomod, sau să căutăm ca, pe lângă valoarea medie, să introducem încă unul sau două numere de același gen pentru descrierea ei, astfel încât acest mic grup de numere să dea în ansamblul său o suficientă caracterizare practică a acelor trăsături ale variabilei studiate, care sunt cele mai esențiale. Să vedem cum poate fi realizat acest lucru.

După cum reiese din exemplele considerate, în multe cazuri problema cea mai importantă din punct de vedere practic este problema abaterii valorilor pe care le ia efectiv variabila întâmplătoare dată, dela valoarea ei medie, adică interesează cât de mult sunt împrăștiate, dispersate aceste valori; vor fi ele, în majoritatea lor, foarte apropiate de valoarea lor medie (deci foarte apropiate ca valoare între ele) sau, dimpotrivă, în majoritatea lor vor diferi simțitor de valoarea medie (și în acest caz unele dintre ele vor diferi, în mod necesar, foarte mult una de alta).

Schema primitivă de mai jos permite să ne reprezentăm în mod clar aceste deosebiri. Să examinăm două variabile întâmplătoare cu următoarea repartiție a probabilităților:

−0,01	+0,01
0,5	0,5

(I)

−100	+100
0,5	0,5

(II)

Ambele variabile întâmplătoare ale căror tabele sunt date mai sus au ca valoare medie 0, dar, pe când prima ia întotdeauna valori foarte apropiate de 0 (și apropiate între ele), a doua variabilă dimpotrivă poate lua numai valori care diferă mult de zero (și între ele). Pentru prima variabilă, cunoașterea valorii ei medii dă și date care ne pot orienta asupra valorilor pe care le poate efectiv lua; pentru a doua mărime, însă, valoarea medie este foarte diferită de valorile efective posibile și nu ne dă nici o ideie despre ele. Spunem că, în al doilea caz, valorile posibile sunt mult mai dispersate decât în primul caz.

Astfel, noi trebue să găsim un număr care să reprezinte în mod rațional o *măsură* a dispersiei unei variabile întâmplătoare, care să ne dea date de orientare, adică să ne indice la ce abateri dela valoarea medie ne putem aștepta la această variabilă. Abaterea $x-\bar{x}$ a variabilei întâmplătoare dela valoarea ei medie $\bar{x}$ este evident și ea o variabilă întâmplătoare; valoarea absolută

Scanned with OKEN Scanner

$|x-\bar{x}|$ a acestei abateri va fi și ea o variabilă întâmplătoare caracterizând mărimea ei, independent de semn. Este de dorit să obținem un număr care să poată caracteriza această abatere întâmplătoare $|x-\bar{x}|$, care să ne arate cât de mare, de exemplu, ar putea fi această abatere. Pentru rezolvarea acestei probleme există mai multe metode, dintre care cele mai folosite în practică sunt următoarele trei:

1. *Abaterea medie.* Cel mai firesc este să luăm valoarea medie $\overline{|x-\bar{x}|}$ ca valoare de orientare pentru variabila întâmplătoare $|x-\bar{x}|$. Această valoare medie a valorii absolute a abaterii se numește *abaterea medie* a variabilei x. Dacă variabila întâmplătoare x este dată în tabela

x_1	x_2	...	x_k
p_1	p_2	...	p_k

tabela variabilei întâmplătoare $|x-\bar{x}|$ are forma:

$\lvert x_1-\bar{x}\rvert$	$\lvert x_2-\bar{x}\rvert$	...	$\lvert x_k-\bar{x}\rvert$
p_1	p_2	...	p_k

unde $\bar{x}=\sum_{i=1}^{k} x_i p_i$; pentru abaterea medie M_x a variabilei x obținem formula

$$M_x=\overline{|x-\bar{x}|}=\sum_{i=1}^{k}|x_i-\bar{x}|\,p_i,$$

în care avem din nou $\bar{x}=\sum_{i=1}^{k} x_i p_i$. Pentru variabilele date în tabelele (I) și (II), $\bar{x}=0$ și obținem respectiv

$$M_{x_I}=0{,}01 \text{ și } M_{x_{II}}=100.$$

De altfel, amândouă aceste exemple sunt banale, deoarece în ambele cazuri valoarea absolută a abaterii nu poate lua decât o singură valoare, pierzând astfel caracterul de variabilă întâmplătoare.

Să mai calculăm abaterile medii ale variabilelor întâmplătoare determinate de tabelele (I') și (II'') din pag. 62. Am văzut

Scanned with OKEN Scanner

acolo că valorile medii ale acestor variabile sunt respectiv 2,1 și 2,2 adică foarte apropiate una de alta. Abaterea medie va fi pentru prima mărime:

$$|1-2,1|\times 0,4+|2-2,1|\times 0,1+|3-2,1|\times 0,5=0,9,$$

iar pentru a doua:

$$|1-2,2|\times 0,1+|2-2,2|\times 0,6+|3-2,2|\times 0,3=0,48.$$

Vedem că pentru a doua mărime abaterea medie este aproape de două ori mai mică decât pentru prima; din punct de vedere practic aceasta înseamnă, evident, că cu toate că în medie amândoi trăgătorii realizează cam același număr de puncte și pot fi consideraţi în acest sens ca fiind la fel de iscusiți, la al doilea trăgător tragerea are un caracter mai omogen, rezultatele sale sunt mult mai puțin dispersate decât la primul trăgător, care pentru același număr mediu de puncte trage inegal, dând deseori rezultate mult mai bune sau mult mai slabe decât cele medii.

2. *Abaterea medie pătratică.* Este foarte natural să măsurăm valoarea de orientare a abaterii cu ajutorul abaterii medii, dar totodată foarte incomod din punct de vedere practic, întrucât calculele și evaluările cu mărimi absolute sunt deseori complicate, iar uneori chiar cu totul inaccesibile. Deaceea se preferă în practică să se introducă o altă măsură pentru evaluarea abaterii.

La fel ca și abaterile $x-\bar{x}$ ale variabilei întâmplătoare x dela valoarea ei medie $\bar{x}$, pătratul $(x-\bar{x})^2$ al acestei abateri este o variabilă întâmplătoare a cărei tabelă, păstrând vechile notații, este:

$(x_1-\bar{x})^2$	$(x_2-\bar{x})^2$	. . .	$(x_k-\bar{x})^2$
p_1	p_2	. . .	p_k

valoarea medie a acestei variabile este deci

$$\sum_{i=1}^{k}(x_i-\bar{x})^2 p_i.$$

Această mărime ne dă o ideie asupra valorii aproximative a *pătratului* abaterii $x-\bar{x}$; deaceea, extrăgând rădăcina pătrată din această mărime

$$Q_x=\sqrt{\sum_{i=1}^{k}(x_i-\bar{x})^2 p_i},$$

Scanned with OKEN Scanner

obținem o mărime ce poate caracteriza măsura aproximativă a abaterii însăși. Mărimea Q_x se numește *abatere medie pătratică* a variabilei întâmplătoare x, iar pătratul ei, adică mărimea Q_x^2 — *dispersia* ei. Desigur această nouă măsură a mărimii abaterii are un caracter ceva mai artificial decât abaterea medie pe care am introdus-o mai înainte: aici mergem pe o cale ocolită, aflând întâi valoarea de orientare a *pătratului* abaterii și abia după aceea ne întoarcem la abaterea însăși, extrăgând rădăcina pătrată. In schimb, după cum vom vedea în paragraful următor, folosirea abaterilor medii pătratice Q_x simplifică simțitor calculele. Din această cauză, statisticienii folosesc de preferință în practica lor abaterile medii pătratice.

Exemplu. Pentru variabilele întâmplătoare determinate de tabelele (I') și (II') pag. 62 avem respectiv

$$Q^2_{x_{I'}} = (1-2{,}1)^2 \times 0{,}4 + (2-2{,}1)^2 \times 0{,}1 + (3-2{,}1)^2 \times 0{,}5 = 0{,}89,$$

și

$$Q^2_{x_{II'}} = (1-2{,}2)^2 \times 0{,}1 + (2-2{,}2)^2 \times 0{,}6 + (3-2{,}2)^2 \times 0{,}3 = 0{,}36,$$

și, deci,

$$Q_{x_{I'}} = \sqrt{0{,}89} \approx 0{,}94 \text{ și } Q_{x_{II'}} = 0{,}6.$$

Pentru aceleași variabile am avut mai sus abaterile medii:

$$M_{x_{I'}} = 0{,}9 \text{ și } M_{x_{II'}} = 0{,}48.$$

Vedem că abaterea medie pătratică, ca și abaterea medie, este mult mai mare pentru prima variabilă decât pentru a doua; măsurând dispersia fie cu abaterea medie, fie cu abaterea medie pătratică, ajungem în ambele cazuri la aceeași concluzie și anume, că prima dintre cele două variabile este dispersată într'o măsură mai mare decât a doua.

In ambele cazuri, abaterea medie pătratică a ieșit mai mare decât abaterea medie; este ușor să ne dăm seama că acest lucru se întâmplă pentru orice variabilă întâmplătoare. In adevăr, dispersia Q_x^2 ca valoare medie a pătratului variabilei $(x-\bar{x})$, conform regulii demonstrate în pag. 67, nu poate fi mai mică decât pătratul valorii medii M_x a variabilei $(x-\bar{x})$, iar din $Q_x^2 \geq M_x$ rezultă $Q_x \geq M_x$.

3. *Abaterea mediană (probabilă).* Adeseori, mai ales în știința militară, este folosită o altă metodă pentru determinarea măsurii dispersiei; o vom expune în termenii unui exemplu de artilerie.

Să presupunem că se trage un foc de artilerie din punctul O pe direcția OX (fig. 9); distanța x dintre locul de cădere a proec

Scanned with OKEN Scanner

tilului și locul din care a pornit este o variabilă întâmplătoare a cărei valoare medie ne indică poziția „centrului de nimeriri" $C(OC=\bar{x})$ în jurul căruia se dispersează mai mult sau mai puțin punctele de cădere a diferitelor proectile.

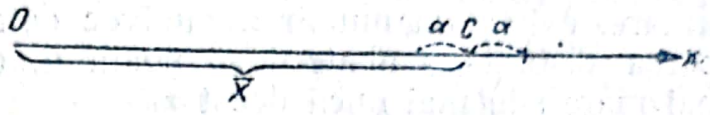

Fig. 9.

Abaterea $x-\bar{x}$ a variabilei întâmplătoare pe care o studiem (lungimea de sbor a proectilului) dela valoarea ei medie este totodată abaterea punctului de cădere a proectilului dela centrul de nimeriri C; deci orice evaluare a variabilei $|x-\bar{x}|$ evaluează totodată gradul de dispersie, de împrăștiere a proectilelor în jurul acestui centru C și servește astfel drept cel mai important indiciu al calității tragerii.

Dacă vom măsura la stânga și la dreapta punctului C câte un segment mic de lungime α, doar puține proectile vor cădea înăuntrul intervalului de lungime 2α cu centrul în punctul C, cu alte cuvinte probabilitatea ca $|x-\bar{x}|<\alpha$ va fi încă foarte mică, când α este mic. Vom lărgi acum intervalul pe care l-am construit, mărind numărul α (care a fost ales arbitrar). Cu cât intervalul construit va fi mai mare, cu atât mai multe proectile vor cădea înăuntrul lui și cu atât mai mare va fi probabilitatea de a nimeri înăuntrul acestui interval pentru un proectil izolat; când α este foarte mare, în practică toate proectilele vor cădea înăuntrul acestui interval; astfel, pe măsură ce α crește, probabilitatea inegalității

$$|x-\bar{x}|<\alpha$$

crește dela zero la unu; la început, când α este mic, este mai probabil ca

$$|x-\bar{x}|>\alpha,$$

adică este mai probabil ca proectilul să cadă în afara intervalului; apoi, când α crește, este mai probabil ca $|x-\bar{x}|<\alpha$, adică ca proectilul să cadă înăuntrul intervalului. Deaceea la trecerea dela valori mici ale numărului α la valori mai mari, trebue să existe o valoare α, a acestui număr astfel încât proectilul să aibă aceeași probabilitate de a cădea înăuntrul sau în afara intervalului de lungime $2\alpha_0$, cu centrul în C. Cu alte cuvinte, inegalitățile

$$|x-\bar{x}|<\alpha_0$$

și

$$|x-\bar{x}|>\alpha_0$$

sunt egal probabile, deci probabilitatea fiecăreia dintre ele este $\frac{1}{2}$ (dacă convenim să neglijem probabilitatea infinit mică a *egalității* exacte $|x-x|=\alpha$). In cazul $\alpha<\alpha_0$, este mai probabilă a doua inegalitate, iar în cazul $\alpha>\alpha_0$, prima inegalitate. Prin urmare, există un număr α_0 univoc determinat, astfel încât valoarea absolută a abaterii să poată fi, cu aceeași probabilitate, mai mare sau mai mică decât α_0.

Valoarea lui α_0 depinde de calitățile tunului cu care se trage: este ușor de văzut că mărimea α_0 poate servi ca măsură a dispersiei proectilelor, ca și abaterea medie pătratică sau abaterea medie. In adevăr, dacă de exemplu α_0 este foarte mic, aceasta înseamnă că jumătate din proectilele trase cad pe un interval foarte mic care înconjoară punctul C, ceeace ne arată că avem o dispersie relativ neinsemnată. Dimpotrivă, dacă α_0 este mare, chiar dacă ne alegem un interval mai mare cu centrul în punctul C, va trebui totuși să ne așteptăm ca jumătate din proectile să cadă în afara acestui interval; aceasta arată, desigur, că obuzele sunt foarte împrăștiate în jurul centrului.

Numărul α_0 se numește deobicei *abaterea mediană sau probabilă* a variabilei x; deci numim abatere mediană sau probabilă a unei variabile întâmplătoare x un număr, astfel determinat încât abaterea $x-x$ în valoare absolută să fie, cu aceeași probabilitate, mai mare sau mai mică decât acest număr. Cu toate că abaterea mediană a variabilei x, pe care o vom însemna prin E_x, nu este mult mai comodă în calculele matematice decât abaterea medie M_x și este mult mai puțin comodă decât abaterea medie pătratică Q_x, în artilerie se obișnuește totuși să se folosească tocmai această mărime E_x pentru aprecierea tuturor abaterilor. Vom vedea în cele ce urmează dece aceasta nu produce deobicei în practică nicio dificultate.

§ 25. Teoreme despre abaterea medie pătratică

Vom arăta acum că abaterile medii pătratice posedă în adevăr însușiri speciale care le fac preferabile tuturor celorlalte măsuri ale abaterii — adică abaterilor medii, mediane (probabile) etc. După cum ne vom convinge ceva mai departe, următoarea problemă are o însemnătate fundamentală în aplicații.

Să presupunem că avem câteva variabile întâmplătoare $x_1, x_2, \ldots, x_n$ cu abaterile medii pătratice $q_1, q_2, \ldots, q_n$. Să considerăm suma $x_1+x_2+\ldots+x_n=X$ și să ne punem problema cum

se află abaterea medie pătratică Q a mărimii X, dacă $q_1, q_2, \ldots, q_n$, sunt date și dacă presupunem variabilele întâmplătoare x_i $(1 \leq i \leq n)$ mutual independente.

Pe baza teoremei adunării valorilor medii avem

$$\bar{X} = \bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_n,$$

deci

$$X - \bar{X} = (x_1 - \bar{x}_1) + (x_2 - \bar{x}_2) + \ldots + (x_n - \bar{x}_n),$$

de unde

$$(X-\bar{X})^2 = \left[\sum_{i=1}^{n}(x_i-\bar{x}_i)\right]^2 = \sum_{i=1}^{n}(x_i-\bar{x}_i)^2 + \sum_{\substack{i=1 \\ i\neq k}}^{n}\sum_{k=1}^{n}(x_i-\bar{x}_i)(x_k-\bar{x}_k). \quad (1)$$

Observăm acum că

$$\overline{(X-\bar{X})^2} = Q^2, \quad \overline{(x_i-\bar{x}_i)^2} = q_i^2 \qquad (1 \leq i \leq n);$$

folosind regula de adunare a valorilor medii obținem

$$Q^2 = \sum_{i=1}^{n} q_i^2 + \sum_{\substack{i=1 \\ i\neq k}}^{n}\sum_{k=1}^{n}\overline{(x_i-\bar{x}_i)\cdot(x_k-\bar{x}_k)}. \quad (2)$$

Dar, întrucât conform ipotezei variabilele x_i și x_k sunt mutual independente $(i \neq k)$ regula înmulțirii valorilor medii ale unor variabile mutual independente ne dă pentru $i \neq k$.

$$\overline{(x_i-\bar{x}_i)\cdot(x_k-\bar{x}_k)} = \overline{(x_i-\bar{x}_i)}\cdot\overline{(x_k-\bar{x}_k)}.$$

Aici, ambii factori din membrul doi sunt nuli pentrucă, de exemplu,

$$\overline{(x_i-\bar{x}_i)} = \bar{x}_i - \bar{x}_i = 0;$$

așa dar, în ultima sumă a egalității (2) fiecare termen în parte este nul ceeace ne duce la relația

$$Q^2 = \sum_{i=1}^{n} q_i^2,$$

adică, *dispersia sumei unor variabile întâmplătoare mutual independente este egală cu suma dispersiilor lor.*

Vedem astfel că, în cazul variabilelor întâmplătoare mutual independente, la regula adunării valorilor medii se mai adaugă o foarte importantă *regulă de adunare a dispersiilor;* pentru abaterile medii pătratice rezultă de aici

$$Q = \sqrt{\sum_{i=1}^{n} q_i^2}. \quad (3)$$

Această posibilitate de a exprima în mod simplu abaterea medie pătratică a sumei prin abaterile medii pătratice ale termenilor în

Scanned with OKEN Scanner

cazul independenței lor mutuale, constitue unul din principalele avantaje ale abaterilor medii pătratice, față de abaterile medii mediane și altele.

Exemplul 1. La n focuri cu aceeași probabilitate de nimerire p, valoarea medie a numărului nimeririlor este np (după cum am văzut la pagina 70). Pentru a aprecia cel puțin aproximativ cât de mare ar putea fi abaterea numărului de trageri reușite dela această valoare medie, vom calcula abaterea medie pătratică a numărului de nimeriri; aceasta se face cel mai simplu prin aplicarea formulei (3).

In adevăr, numărul tragerilor reușite din n focuri poate fi privit (așa cum am făcut la pagina 70), ca suma numerelor de nimeriri ale focurilor separate și, deoarece considerăm aceste numere ca variabile întâmplătoare mutual independente, conform regulii de adunare a dispersiilor, la calculul abaterii medii pătratice Q a numărului total de nimeriri putem folosi formula (3) unde $q_1, q_2, \ldots, q_n$ reprezintă abaterile medii pătratice ale numerelor de nimeriri la focuri separate. Numărul de nimeriri x_i la cel de al i-lea foc se determină însă din tabloul

1	0
p	$1-p$

;

deci

$$\bar{x}_i = p \text{ și } q_i^2 = \overline{(x-\bar{x}_i)^2} = (1-p)^2 p + p^2(1-p) = p(1-p);$$

deci

$$Q = \sqrt{\sum_{i=1}^{n} q_i^2} = \sqrt{np(1-p)}.$$

Problema este astfel rezolvată. Comparând valoarea medie np a numărului de nimeriri cu abaterea sa medie pătratică $\sqrt{np(1-p)}$, vedem că pentru valori mari ale lui n (adică pentru un număr mare de focuri trase) aceasta din urmă este mult mai mică decât prima și reprezintă numai o mică parte din ea. Astfel, pentru $n=900$ și $p=\frac{1}{2}$, valoarea medie a numărului de nimeriri este 450, iar abaterea medie pătratică $\sqrt{900 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 15$, astfel că numărul de trageri reușite va diferi cam cu 3—4% de valoarea sa medie.

Exemplul 2. Să ne închipuim că se montează un mecanism oarecare, compus din n detalii, care se îmbină dealungul unei axe

Scanned with OKEN Scanner

oarecare și sunt încadrate la extremități de o piesă (v. fig. 10). Lungimea fiecărei piese poate diferi întrucâtva de standardul corespunzător și este deci o variabilă întâmplătoare. Vom presupune că aceste variabile întâmplătoare sunt mutual independente. Dacă lungimile medii ale pieselor și dispersiile acestor lungimi sunt respectiv $a_1, a_2, \ldots a_n$ și $q_1, q_2, \ldots, q_n$, valoarea medie și dispersa lungimii lanțului de n piese vor fi respectiv

Fig. 10

$$a=\sum_{k=1}^{n} ak \text{ și } q=\sqrt{\sum_{k=1}^{n} q_k^2}.$$

In particular, dacă $n = 9$, $a_1 = a_2 = \ldots = a_9 = 10$ cm și $q_1 = q_2 = \ldots = q_9 = 0{,}2$ cm, atunci $a = 90$ cm și $q = \sqrt{9 \cdot (0{,}2)^2} = 0{,}6$ cm.

Vedem astfel că, dacă în medie lungimea fiecărei piese în parte se abate dela valoarea ei medie cu 2%, lungimea lanțului format cu aceste detalii diferă de valoarea sa medie numai cu circa $\frac{2}{3}$ %. Acest fapt — scăderea erorii relative la adunarea variabilelor întâmplătoare — are un rol însemnat la montarea mecanismelor de precizie. In adevăr, dacă n'ar exista o compensare reciprocă între abaterile dimensiunilor diverselor piese față de dimensiunile normale cerute, la montarea mecanismelor s'ar ivi întotdeauna cazuri când detaliile care încadrează mecanismul la extremități nu ar cuprinde lanțul de piese sau când, dimpotrivă, ar rămâne spații libere prea mari. In ambele cazuri am avea rebuturi; nu ar fi indicat să se lupte contra acestor rebuturi diminuând „toleranțele", adică micșorând abaterile admisibile ale dimensiunilor efective ale pieselor față de cele stabilite, întrucât o sporire relativ neînsemnată a preciziei prelucrării ar ridica costul mecanismului [1].

Exemplul 3. Să presupunem că în condiții invariabile se fac n măsurători ale unei mărimi oarecare. Datorită unei serii întregi de circumstanțe (poziția aparatului, a observatorului, variații în starea aerului, prezența prafului, etc.) diferitele măsurători vor da, în general, rezultate diferite — au loc erori accidentale ale măsurătorilor. Vom însemna rezultatele măsurărilor prin $x_1, x_2, \ldots, x_n$, adăugând fiecărui x ca indice numărul de ordine al măsurătorii. Valoarea medie pentru toate aceste variabile întâmplătoare este aceeași —x. Evident, este firesc să presupunem

[1] Gândirea tehnică a ajuns în ultimul timp la concluzia că este necesar să se creeze o teorie a toleranțelor, bazată pe considerente de calculul probabilităților. Această teorie a toleranțelor este elaborată în prezent de către savanții sovietici.

că și abaterea pătratică medie q este una și aceeași în toate măsurătorile, întrucât ele se fac în condiții invariabile. În sfârșit, vom presupune, ca deobicei, că variabilele $x_1, x_2, \ldots, x_n$ sunt mutual independente.

Să considerăm acum media aritmetică

$$\xi = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

a rezultatelor celor n măsurători. Aceasta este o variabilă întâmplătoare; să căutăm valoarea ei medie și abaterea ei medie pătratică. Conform regulii de adunare,

$$\bar{\xi} = \frac{1}{n}(\overline{x_1 + x_2 + \ldots + x_n}) = \frac{1}{n}(\bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_n) = \frac{1}{n}(n\bar{x}) = \bar{x},$$

adică valoarea medie este aceeași ca la fiecare măsurătoare în parte, ceeace de altfel era clar dela bun început.

Apoi, conform regulii de adunare a dispersiilor (3), abaterea medie pătratică a sumei $x_1 + x_2 + \ldots + x_n$, este

$$Q = \sqrt{nq^2} = q\sqrt{n},$$

iar abaterea medie pătratică a mărimii ξ, care reprezintă $\frac{1}{n}$ din această sumă, este

$$\frac{Q}{n} = \frac{q}{\sqrt{n}}.$$

Am ajuns astfel la un rezultat foarte important: *media aritmetică a n variabile întâmplătoare mutual independente cu repartiții identice are:*

a) *aceeași valoare medie ca fiecare din variabilele componente;*

b) *o abatere medie pătratică de $\sqrt{n}$ ori mai mică decât fiecare din variabilele componente.*

Astfel, dacă valoarea medie a mărimii măsurate este $\bar{x} = 200$ m, iar abaterea medie pătratică $q = 5$ m, media aritmetică ξ a rezultatelor unei sute de măsurători va avea, evident, ca valoare medie același număr 200 m; însă abaterea sa medie pătratică va fi de $\sqrt{100} = 10$ ori mai mică decât la o măsurătoare separată, adică va fi numai $\frac{q}{10} = 0{,}5$ m. Astfel, este firesc să ne așteptăm ca media aritmetică a rezultatelor unei sute de măsurători să fie mult mai apropiată de valoarea ei medie 200 m, decât rezultatul uneia dintre măsurători. *Media aritmetică a unui mare număr de variabile mutual independente are o dispersie mult mai mică decât fiecare variabilă în parte.*

CAPITOLUL XI

LEGEA NUMERELOR MARI

§ 26. Inegalitatea lui Cebâșev

Am mai arătat până acum că cunoașterea uneia dintre abaterile medii ale unei variabile întâmplătoare (de exemplu, a abaterii medii pătratice) ne permite să ne facem o idee aproximativă despre mărimea la care trebue să ne așteptăm pentru abaterile valorilor luate efectiv de această variabilă dela valoarea ei medie. Totuși această constatare nu conține încă prin ea însăși nici o apreciere cantitativă și nu permite să calculăm nici măcar aproximativ probabilitatea abaterilor mari. Toate acestea ne conduc la următorul raționament simplu, pe care cel dintâi l-a făcut Cebâșev. Vom pleca dela expresia dispersiei variabilei întâmplătoare x (pag. 77):

$$Q_x^2 = \sum_{i=1}^{k} (x_i - \bar{x})^2 p_i.$$

Presupunem că α este un număr pozitiv oarecare; dacă din suma de mai sus eliminăm toți termenii pentru care $|x - \bar{x}| \leq \alpha$, și lăsăm numai termenii pentru care $|x - \bar{x}| > \alpha$, suma poate numai să se micșoreze, adică

$$Q_x^2 \geq \sum_{|x_i - \bar{x}| > \alpha} (x_i - \bar{x})^2 p_i.$$

Această sumă se va micșora și mai mult dacă în fiecare termen al ei vom înlocui factorul $(x_i - \bar{x})^2$ prin mărimea mai mică α^2:

$$Q_x^2 \geq \alpha^2 \sum_{|x_i - \bar{x}| > \alpha} p_i.$$

Suma din partea dreaptă reprezintă acum suma probabilităților tuturor acelor valori x_i ale variabilei întâmplătoare x, care se abat dela $\bar{x}$ de o parte și de alta cu mai mult decât α; conform regulii de adunare, aceasta este probabilitatea ca mărimea x să ia una dintre aceste valori. Cu alte cuvinte, aceasta este probabilitatea $P(|x - \bar{x}| > \alpha)$ ca abaterea reală să fie mai mare decât α; obținem astfel

$$P(|x - \bar{x}| > \alpha) \leq \frac{Q_x^2}{\alpha^2}, \qquad (1)$$

ceeace nu permite să apreciem probabilitatea abaterilor mai mari decât un număr α dat dinainte, cu condiția numai să fie cunoscută abaterea medie pătratică Q_x. Este adevărat că evaluarea dată de „inegalitatea lui Cebâșev" (1) este adeseori prea grosolană, dar

Scanned with OKEN Scanner

ea poate fi totuși folosită câteodată nemijlocit în practică, fără să mai vorbim despre importanța ei teoretică foarte mare.

La sfârșitul paragrafului precedent, am examinat următorul exemplu: valoarea medie a rezultatelor măsurătorii este 200 m, iar abaterea medie pătratică 5 m; în aceste condiții, probabilitatea de a obține efectiv o abatere mai mare de 3 m este foarte posibilă (este de bănuit că ea este mai mare decât jumătate; valoarea ei exactă poate fi desigur aflată numai dacă se cunoaște perfect legea de repartiție a rezultatelor măsurătorilor). Am văzut însă că, pentru media aritmetică ξ a rezultatelor unei sute de măsurători, abaterea medie pătratică este de numai 0,5 m. Deaceea, în virtutea inegalității (1)

$$P(|\xi-200|>3) \leq \frac{(0,5)^2}{3^2}=\frac{1}{36}\approx 0,03.$$

Astfel, pentru media aritmetică a 100 de măsurători, probabilitatea de a obține o abatere de peste 3 m este foarte mică (în realitate ea este mult mai mică chiar decât limita obținută, astfel încât probabilitatea unei asemenea abateri este, practic, cu totul neglijabilă).

In exemplul 1 dela pagina 82—83 am avut, pentru numărul de nimeriri din 900 focuri, valoarea medie 450 și abaterea medie pătratică 15; pentru probabilitatea ca numărul efectiv m de trageri reușite să fie cuprins, de exemplu, între 400 și 500 (adică $|m-450| \leq 50$), rezultă din inegalitatea lui Cebâșev:

$$P(|m-450| \leq 50)=1-P(|m-450|>50) \geq 1-\frac{15^2}{50^2}=0,91.$$

In realitate această probabilitate este însă mult mai mare.

§ 27. Legea numerelor mari

Să considerăm n variabile întâmplătoare mutual independente $x_1, x_2, \ldots, x_n$ cu aceeași valoare medie a și cu aceeași abatere medie pătratică q; pentru media aritmetică a acestor variabile

$$\xi=\frac{x_1+x_2+\ldots+x_n}{n},$$

după cum am văzut la pag. 84, valoarea medie este a, iar abaterea medie pătratică este $\frac{q}{\sqrt{n}}$; deci inegalitatea lui Cebâșev dă pentru un α pozitiv oarecare,

$$P(|\xi-a|>\alpha) \leq \frac{q^2}{\alpha^2 n}. \qquad (2)$$

Scanned with OKEN Scanner

Să presupunem, de exemplu, că este vorba de media aritmetică a n măsurători ale unei mărimi oarecare și fie, ca mai înainte, $\varrho=5$ m, $a=200$ m. Obținem atunci

$$P(|\xi-200|>\alpha) \leq \frac{25}{\alpha^2 n}.$$

Putem alege α foarte mic, de exemplu $\alpha=0{,}5$ m; atunci

$$P(|\xi-200|>0{,}5) \leq \frac{100}{n}.$$

Dacă alegem numărul de măsurători n suficient de mare, membrul din dreapta al acestei inegalități poate fi făcut oricât de mic; astfel, pentru $n=10\,000$ obținem

$$P(|\xi-200|>0{,}5) \leq 0{,}01.$$

ξ fiind media aritmetică a 10 000 măsurători. Dacă neglijăm posibilitatea unor evenimente atât de puțin probabile, putem spune că la 10 000 măsurători, media lor aritmetică va diferi cu siguranță cu cel mult 50 cm în plus sau în minus de 200 m. Dacă dorim să atingem o apropiere și mai mare, de exemplu 10 cm, trebue să punem $\alpha=0{,}1$ m și obținem

$$P(|\xi-200|>0{,}1) \leq \frac{25}{0{,}01n} = \frac{2\,500}{n}.$$

Pentru a face membrul din dreapta al acestei inegalități mai mic decât 0,01, ar trebui să luăm un număr de măsurători egal nu cu 10 000 (ceeace n'ar mai fi suficient în acest caz), ci cu 250 000. Este evident că, în general, oricât de mic ar fi α, putem face membrul drept al inegalității (2) oricât de mic, cu condiția să luăm un n suficient de mare. Astfel, pentru un n suficient de mare putem considera inegalitatea inversă

$$|\xi-a| \leq \alpha$$

oricât de apropiată de certitudine.

Dacă variabilele întâmplătoare $x_1, x_2, \ldots, x_n$ sunt mutual independente și dacă au toate aceeași valoare medie a și aceeași abatere medie pătratică, atunci pentru un n suficient de mare, mărimea

$$\xi=\frac{x_1+x_2+\ldots+x_n}{n}$$

va diferi oricât de puțin de a cu o probabilitate oricât de apropiată de unitate (practic cu certitudine).

Acesta este un caz particular simplu al unela din principalele

teoreme ale calculului probabilităţilor, aşa numita *lege a numerelor mari*, descoperită pe la mijlocul secolului trecut de marele matematician rus Cebâşev. Conţinutul profund al acestei legi remarcabile constă în aceea că, pe când o variabilă întâmplătoare izolată poate (după cum ştim) să ia adesea valori depărtate de valoarea ei medie (adică să aibă o împrăştiere apreciabilă), media aritmetică a unui mare număr de variabile întâmplătoare se comportă sub acest raport cu totul altfel: o astfel de mărime are o împrăştiere foarte mică şi cu o probabilitate covârşitoare ia numai valori foarte apropiate de valoarea ei medie. Aceasta se explică, desigur, prin faptul că, luându-se media aritmetică, abaterile întâmplătoare într'o parte şi alta se anihilează reciproc, astfel încât în majoritatea cazurilor abaterea totală este mică.

O aplicaţie foarte importantă şi întâlnită deseori în practică a teoremei lui Cebâşev demonstrate aici constă în aprecierea calităţii unei mari cantităţi dintr'un material omogen, printr'o probă relativ mică (sondaj). Astfel, de exemplu, putem aprecia calitatea bumbacului aflat într'un balot, judecând după câteva mănunchiuri mici, extrase la întâmplare din diferite puncte ale balotului. Tot aşa calitatea unei partide mari de grăunţe se judecă după câteva probe scoase la întâmplare cu o sondă, din diferite locuri ale partidei în chestiune [1].

Constatările privitoare la calitatea producţiei, făcute pe baza unui asemenea sondaj, prezintă o mare precizie întrucât cantitatea grăunţelor cuprinse în sondă, mică în comparaţie cu întreaga cantitate de grăunţe, este mare prin ea însăşi şi permite, conform legii numerelor mari, să apreciem cu o precizie suficientă greutatea medie a unui bob şi prin urmare, să apreciem calitatea întregii cantităţi de grăunţe. Tot astfel, putem judeca despre un balot de bumbac de 20 puduri după o cantitate mică (o probă) care conţine câteva sute de fibre cântărind în total abia a zecea parte dintr'un gram.

§ 28. Demonstraţia legii numerelor mari

Până acum am considerat numai cazul în care toate variabilele $x_1, x_2, \ldots$ au una şi aceeaşi valoare medie şi aceeaşi abatere medie pătratică. Legea numerelor mari se aplică însă şi în cazul unor ipoteze mai generale. Vom examina acum cazul în care valorile medii ale variabilelor $x_1, x_2, \ldots$ sunt numere oarecare (pe care

[1] Sonda extrage circa 100—200 g odată, iar întreaga partidă conţine zeci şi poate chiar sute de tone de grăunţe.

le vom insemna respectiv prin $a_1, a_2, \ldots$), în general diferite. În acest caz, valoarea medie a mărimii

$$\xi=\frac{1}{n}(x_1+x_2+\ldots+x_n),$$

va fi

$$A=\frac{1}{n}(a_1+a_2+\ldots+a_n),$$

și în virtutea inegalității lui Cebâșev (1), pentru un α pozitiv oarecare avem

$$P(|\xi-A|>\alpha)\leq\frac{Q_\xi^2}{\alpha^2}. \tag{3}$$

După cum se vede, totul se reduce la evaluarea mărimii Q_ξ^2, care este aproape tot atât de simplă aici, ca și în cazul particular studiat mai înainte. Q_ξ^2 este dispersia mărimii ξ egală cu suma a n variabile întâmplătoare mutual independente împărțite cu n (păstrăm, bineînțeles, ipoteza independenței mutuale). Conform regulii de adunare a dispersiilor obținem

$$Q_\xi^2=\frac{1}{n^2}(q_1^2+q_2^2+\ldots+q_n^2),$$

unde $q_1, q_2, \ldots$ reprezintă respectiv abaterile medii pătratice ale variabilelor $x_1, x_2, \ldots$ În cazul general, aceste abateri medii pătratice diferă între ele. Vom admite însă că oricât de multe variabile am lua (adică oricât de mare ar fi numărul n), abaterile medii pătratice ale tuturor acestor variabile vor fi totdeauna mai mici decât un anumit număr pozitiv b. În practică, această condiție este totdeauna îndeplinită, întrucât avem de adunat variabile mai mult sau mai puțin de același tip și dispersia lor nu diferă prea mult la diferitele variabile. Astfel, vom admite că $q_i<b\,(i=1, 2, \ldots)$; în acest caz din ultima egalitate rezultă însă

$$Q_\xi^2<\frac{1}{n^2}nb^2=\frac{b^2}{n},$$

astfel încât din inegalitatea (3) deducem

$$P(|\xi-A|>\alpha)<\frac{b^2}{n\alpha^2}.$$

Oricât de mic ar fi α pentru un număr n destul de mare de variabile întâmplătoare membrul drept al acestei inegalități poate fi făcut oricât de mic. Cu aceasta legea numerelor mari este demonstrată în cazul general pe care îl studiem.

Astfel, *dacă variabilele $x_1, x_2, \ldots$ sunt mutual independente, iar abaterile lor medii pătratice rămân toate mai mici decât un*

anumit număr pozitiv, ne putem aştepta ca, pentru un număr de variabile n suficient de mare, media lor aritmetică

$$\xi=\frac{1}{n}(x_1+x_2+\ldots+x_n)$$

să aibă, cu o probabilitate apropiată de unitate, abateri oricât de mici în valoare absolută.

Aceasta este tocmai legea numerelor mari, în forma generală pe care a dat-o Cebâşev.

Este cazul să subliniem acum o circumstanţă esenţială. Să presupunem că se măsoară o mărime oarecare a. Repetând măsurătorile în condiţii invariabile, observatorul obţine rezultate numerice întrucâtva diferite $x_1, x_2, \ldots, x_n$. Ca valoare aproximativă a lui a, el ia media aritmetică

$$a\sim\frac{1}{n}(x_1+x_2+\ldots+x_n).$$

Putem spera să obţinem o valoare oricât de apropiată de a, dacă efectuăm un număr destul de mare de experienţe?

Da, aceasta se poate întâmpla dacă măsurătorile se fac fără erori sistematice, adică dacă

$$\overline{x}_k=a \quad (\text{pentru } k=1, 2, \ldots, n)$$

şi dacă valorile înseşi nu sunt nedeterminate, cu alte cuvinte dacă la măsurători citim pe aparatul folosit exact indicaţiile care se obţin. Dacă însă aparatul este construit în aşa fel încât nu permite citirea decât cu o precizie sub o anumită valoare δ (de exemplu, pentrucă cea mai mică diviziune a scalei pe care se face citirea este egală cu δ) atunci nu ne putem aştepta firește la o precizie mai mare de $\pm\delta$. Este clar că, în acest caz, media aritmetică va avea aceeaşi nedeterminare δ ca şi fiecare x_k.

Observaţia făcută ne învaţă că, dacă aparatele de măsură dau rezultatele măsurătorilor cu o anumită nedeterminare δ, este inutil să căutăm să obţinem, bazaţi pe legea numerelor mari, valoarea lui a cu o precizie mai mare, iar calculele efectuate în aceste condiţii se transformă într'o distracţie aritmetică, fără sens.

Scanned with OKEN Scanner

CAPITOLUL XII

LEGI NORMALE

§ 29. Modul de a pune problema

Am văzut că, într'un mare număr de fenomene ale naturii, precum și în procesele de producție și în operațiile de luptă, un rol însemnat îl joacă anumite variabile întâmplătoare. Deseori, înainte ca fenomenul, procesul sau operația să fi avut loc, tot ce putem cunoaște despre aceste variabile întâmplătoare sunt legile lor de repartiție, adică o enumerare a tuturor valorilor lor posibile cu indicarea probabilităților respective. Dacă variabila poate lua un număr nelimitat de valori diferite (distanța străbătută de un proectil, erorile de măsură, și altele), este preferabil să se dea nu probabilitățile diferitelor valori ale variabilei, ci probabilitățile unor intervale întregi de valori (de exemplu probabilitatea ca un proectil să străbată 720—740 m, sau 740—760 m, etc.). Această considerație nu schimbă esența lucrărilor: pentru a cunoaște o variabilă întâmplătoare, trebue să obținem o imagine cât mai exactă a legii ei de repartiție.

Dacă noi, căutând să aflăm legile de repartiție ale variabilelor întâmplătoare pe care le întâlnim, am renunța la orice raționament și la orice presupunere cu caracter general, considerând fiecare variabilă întâmplătoare fără nici o ipoteză prealabilă și încercând să stabilim pe cale pur experimentală toate trăsăturile legii ei de repartiție, ne-am afla în fața unei probleme practic irealizabilă prin dificultățile pe care le-ar ridica. In fiecare caz nou, ar trebui să facem un număr imens de experiențe, pentru a stabili măcar trăsăturile principale ale noii legi de repartiție necunoscute. Deaceea, oamenii de știință au căutat încă de mult să găsească tipuri comune de legi de repartiție, care s'ar putea prevedea, bineînțeles numai dacă acestea au o justificare la care să ne așteptăm dacă nu pentru toate variabilele întâmplătoare, cel puțin pentru clase largi de asemenea variabile întâlnite în practică. Unele dintre aceste tipuri au fost stabilite de mult în teorie, fiind confirmate apoi de experiență. Firește, este cât se poate de avantajos ca, pe baza raționamentelor teoretice și a experienței acumulate, să avem posibilitatea, atunci când întâlnim o variabilă întâmplătoare, să prevedem tipul legii ei de repartiție. Dacă această propunere se dovedește corectă, un număr relativ mic de experiențe sau observații este deobicei suficient pentru a stabili toate trăsăturile necesare ale legii de repartiție căutate.

Scanned with OKEN Scanner

Cercetările teoretice au arătat că, într'un număr de cazuri întâlnite în practică, ne putem aştepta cu suficient temei la legi de repartiţie de un anumit tip determinat, numite *legi normale*. In acest capitol vom vorbi pe scurt despre aceste legi, lăsând la o parte, din cauza caracterului lor complicat, toate demonstraţiile şi formulările exacte.

Printre variabilele întâmplătoare pe care le întâlnim în practică, foarte multe au caracter de „erori accidentale" sau cel puţin pot fi reduse uşor la asemenea „erori". Să presupunem, de exemplu, că studiem distanţa x parcursă de un proectil tras dintr'un tun oarecare. Vom admite, cum este şi firesc, că există o distanţă normală medie x_0, pentru care am reglat dispozitivul de ochire; diferenţa $x-x_0$ reprezintă „eroarea" şi studiul variabilei întâmplătoare x se reduce la studiul nemijlocit al „erorii accidentale" $x-x_0$.

Dar o asemenea eroare care-şi schimbă valoarea dela o tragere la alta depinde, în general, de foarte multe cauze, care acţionează independent una de alta; oscilaţii întâmplătoare ale ţevii tunului, diferenţe inevitabile (deşi foarte mici) în greutatea şi forma proectilelor, schimbări întâmplătoare ale condiţiilor atmosferice, care provoacă modificări în rezistenţa aerului, erori întâmplătoare de ochire (dacă aceasta se reface înainte de fiecare tragere sau înaintea fiecărei serii scurte de trageri). Toate aceste cauze şi multe altele pot provoca erori în distanţa străbătută. Toate aceste erori parţiale reprezintă în fond variabile întâmplătoare independente între ele şi anume astfel încât *acţiunea fiecăreia dintre ele reprezintă numai o parte mică a acţiunii lor globale*, iar eroarea definitivă $x-x_0$ pe care vrem s'o studiem va fi tocmai acţiunea însumată a tuturor acestor erori accidentale, provenite din diferite cauze. Astfel, în exemplul nostru, eroarea care ne interesează este suma unui mare număr de variabile întâmplătoare mutual independente şi este evident că lucrurile se vor prezenta la fel, pentru majoritatea erorilor accidentale pe care le întâlnim în practică. Un raţionament teoretic, pe care nu-l putem reproduce aici, arată că legea de repartiţie a unei variabile întâmplătoare, care este suma unui mare număr de variabile întâmplătoare mutual independente, numai în virtutea acestui fapt, trebue să fie apropiată de o lege de un tip absolut determinat [1], oricare ar fi natura termenilor, *cu singura condiţie ca fiecare să fie mic în comparaţie cu întreaga sumă*.

Acest tip este tocmai tipul legilor normale.

Obţinem astfel posibilitatea de a presupune că o parte foarte însemnată dintre variabilele întâmplătoare întâlnite în practică

[1] Pentru această problemă vezi încheierea.

(toate erorire formate dintr'un mare număr de erori mutual independente) este distribuită aproximativ după legi normale.

Acum va trebui să ne familiarizăm cu trăsăturile fundamentale ale acestor legi.

§ 30. Noțiunea de curbă de repartiție

In § 15, am avut ocazia să reprezentăm grafic legile de repartiție cu ajutorul diagramelor; această metodă este foarte utilă, întrucât ne permite ca, fără a mai examina vreo tabelă, să cuprindem dintr'o singură privire trăsăturile principale ale legii de repartiție studiate. Schema acestei reprezentări grafice este următoarea: pe o dreaptă orizontală se dispun toate valorile posibile ale variabilei întâmplătoare date, începând dela o origine O oarecare, valorile pozitive fiind fixate în dreapta, iar cele negative în stânga ei (fig. 11).

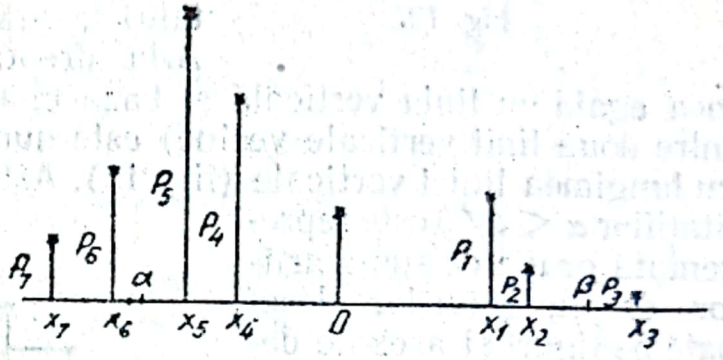

Fig. 11.

In orice punct al dreptei care reprezintă o valoare posibilă a variabilei întâmplătoare ridicăm o verticală, a cărei lungime reprezintă probabilitatea acestei valori. Scara se alege pentru ambele direcții astfel încât întreaga diagramă să aibă o formă comodă și ușor de privit. O privire fugară aruncată asupra figurii 11 ne arată că variabila întâmplătoare pe care ea o caracterizează are ca valoare cea mai probabilă valoarea x_5 (negativă) și, pe măsură ce valorile posibile se îndepărtează de x_5, probabilitățile lor scad necontenit (destul de accentuat). Probabilitatea ca variabila să ia o valoare cuprinsă într'un interval (α, β) oarecare este egală, conform regulii de adunare, cu suma probabilităților tuturor valorilor posibile situate în acest interval și se reprezintă geometric prin suma lungimilor liniilor verticale situate deasupra intervalului în figura 11 $P(\alpha < x < \beta) = p_1 + p_2 + p_4 + p_5$. Dacă numărul valorilor posibile este foarte mare, ceeace se întâmplă adesea în practică, pentru a nu întinde prea mult figura se ia pentru direcția orizontală o scară foarte mică, astfel încât valorile posibile sunt așezate foarte aproape una de alta (fig. 12) și vârfurile liniilor verticale trase se confundă, ne apar ca o curbă continuă numită „curba de repartiție" a variabilei întâmplătoare

date. Evident și aici probabilitatea inegalităților $\alpha < x < \beta$ se reprezintă grafic prin suma lungimilor liniilor verticale situate deasupra intervalului (α, β). Să presupunem acum că distanța între două valori posibile vecine este întotdeauna egală cu unitatea; acest lucru se întâmplă, de exemplu, dacă valorile posibile sunt exprimate printr'un șir de numere întregi consecutive, ceeace se poate realiza practic întotdeauna alegând o unitate de măsură suficient de mică. Aria dreptunghiului având înălțimea egală cu linia verticală și baza egală cu unitatea (distanța între două linii verticale vecine) este numeric egală în acest caz cu lungimea liniei verticale (fig. 13). Astfel, probabilitatea inegalităților $\alpha < x < \beta$ este reprezentată grafic de suma ariilor dreptunghiurilor desenate pe figuri și așezate deasupra intervalului (α, β). Dacă însă valorile posibile sunt foarte apropiate una de alta, ca în fig. 12, suma ariilor acestor dreptunghiuri nu va diferi practic de aria domeniului delimitat în jos de segmentul (α, β) în sus de curba de repartiție, iar în părți de liniile verticale duse din punctele α și β (fig. 14)[1]. Astfel, pe o diagramă curbilinie de tipul celei din fig. 14, probabilitatea ca variabila întâmplătoare dată să ia valori într'un anumit interval se exprimă simplu și comod prin suprafața situată deasupra acestui interval și sub curba de repartiție. Dacă legea de repartiție a variabilei date este reprezentată printr'o astfel de diagramă curbilinie, nu mai este nevoie să se tragă pe ea linii verticale, care nu ar fi de nici un folos și care nu ar face decât să încarce desenul; până și problema probabilității diverselor valori pierde aici din actualitate;

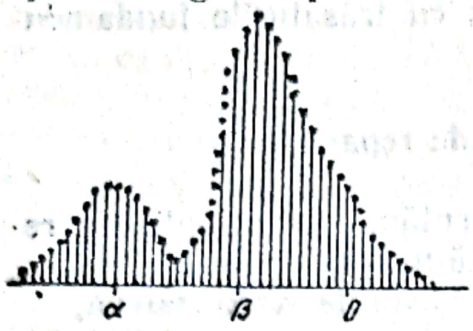

Fig. 12.

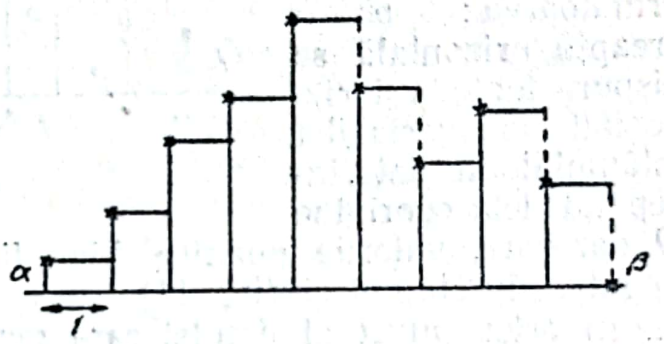

Fig. 13.

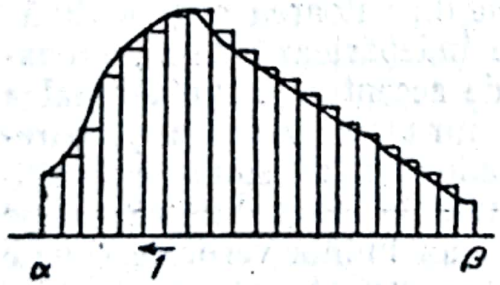

Fig. 14.

[1] In acest caz, considerăm, ca și înainte, ca unitate de lungime, distanța între două valori posibile vecine.

dacă valorile posibile sunt multe la număr (și această ipoteză stă dealtfel la baza tuturor diagramelor curbilinii), atunci probabilitățile diverselor valori vor fi, în mod obișnuit, neglijabile (practic nule) și pierd orice interes. Astfel, la o tragere cu tunul nu este deloc important să se cunoască, de exemplu, probabilitatea ca punctul de cădere al proectilului să se abată dela centrul țintei exact cu 475 cm. Dimpotrivă, un interes esențial prezintă problema probabilității unei abateri cuprinse între 3 și 5 m. La fel stau lucrurile în toate cazurile asemănătoare: dacă o variabilă întâmplătoare ia foarte multe valori, este important să cunoaștem nu probabilitățile diferitelor valori ci probabilitățile unor intervale întregi de valori. Tocmai aceste din urmă probabilități ne sunt însă date direct și intuitiv, după cum am văzut, de ariile din diagramele curbilinii.

§ 31. Proprietățile curbelor normale de repartiție

O variabilă distribuită după o lege normală are întotdeauna o infinitate de valori posibile; de aceea este comod să reprezentăm legile normale prin diagrame curbilinii.

În fig. 15 sunt reprezentate câteva curbe de repartiție după legea normală. Cu toate deosebirile de formă ale acestor curbe, putem observa la ele anumite trăsături comune foarte bine conturate.

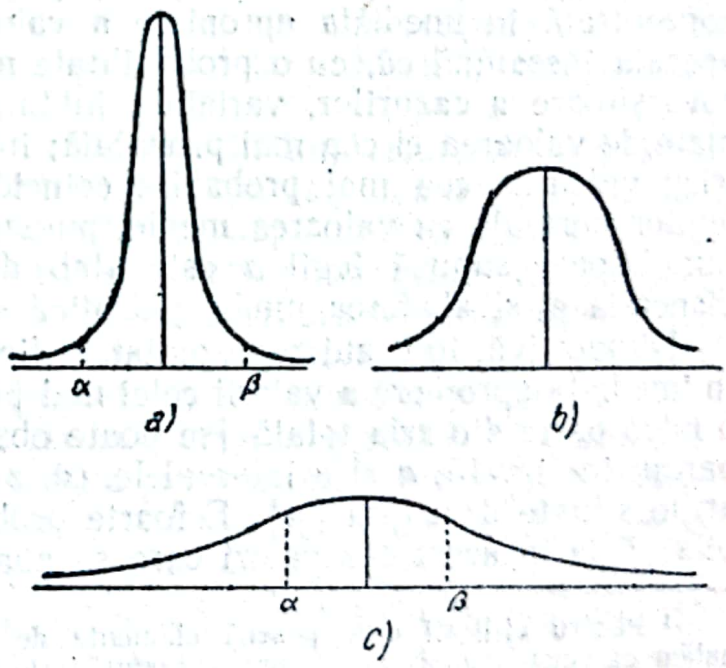

Fig. 15.

1. Toate curbele au un punct de maxim, ele scad necontenit la dreapta și la stânga lui. Aceasta înseamnă evident că pe măsură ce valorile variabilei întâmplătoare se depărtează de valoarea ei cea mai probabilă, probabilitățile lor scad necontenit.

2. Toate curbele sunt simetrice în raport cu o dreaptă verticală dusă prin punctul de maxim. Aceasta înseamnă evident că valorile egal depărtate de valoarea cea mai probabilă au probabilități egale.

Scanned with OKEN Scanner

3. Toate curbele au formă de clopot: în apropierea punctului de maxim ele au convexitatea îndreptată în sus, iar la o anumită distanță de el curbele își schimbă convexitatea; această distanță (ca și înălțimea maximă) diferă dela o curbă la alta [1].

Prin ce se deosebesc una de alta diferitele curbe normale? Pentru a da un răspuns clar la această problemă, trebue să ne amintim mai întâi că, pentru fiecare curbă de repartiție, întreaga arie situată sub ea este egală cu unitatea, deoarece ea trebue să fie egală cu probabilitatea ca variabila întâmplătoare să ia o valoare oarecare din valorile ei posibile, adică este egală cu probabilitatea unui eveniment cert. Deosebirea între curbele de repartiție constă deci numai în faptul că această arie totală, aceeași pentru toate curbele, este distribuită în mod diferit între diferitele intervale. După cum ne arată curbele din fig. 15, pentru legile normale, problema fundamentală este de a determina ce parte din aria totală este concentrată deasupra intervalului din jurul valorii celei mai probabile și ce parte din arie este situată deasupra porțiunilor îndepărtate de valoarea cea mai probabilă. Pentru legea reprezentată în fig. 15, *a*, aproape întreaga arie este concentrată în imediata apropiere a valorii celei mai probabile; aceasta înseamnă că, cu o probabilitate mare deci în majoritatea covârșitoare a cazurilor, variabila întâmplătoare ia valori apropiate de valoarea ei cea mai probabilă; întrucât, în virtutea simetriei, valoarea cea mai probabilă coincide întotdeauna în cazul legilor normale cu valoarea medie, putem spune că variabila întâmplătoare supusă legii *a* este slab dispersată; în particular dispersia ei și abaterea medie pătratică sunt mici.

Dimpotrivă, în cazul reprezentat în fig. 15, *c*, aria concentrată în imediata apropiere a valorii celei mai probabile constitue numai o mică parte din aria totală [se poate observa imediat dacă comparăm în fig. 15, *a* și *c* intervalele (α, β) de aceeași lungime și ariile situate deasupra lor]. E foarte probabil, așa dar, ca variabila să ia în acest caz valori care se abat mult dela valoarea ei

---

[1] Pentru cititorii care posedă elemente de matematici superioare semnalăm că ecuația unei curbe care reprezintă o lege normală are forma

$$y=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

unde numărul $e=2{,}71828\ldots$ este baza logaritmilor naturali, $\pi=3{,}14159\ldots$ este raportul dintre lungimea circomferinței și diametrul ei, iar mărimile a și σ^2 reprezintă respectiv valoarea medie și dispersia variabilei întâmplătoare. Cunoașterea expresiei analitice a legii normale poate ușura mult studiul materialului ce urmează în această carte. Totuși, expunerea a tot ce urmează este accesibilă și cititorului care nu are deloc cunoștințe de matematici superioare.

cea mai probabilă. Variabila este puternic dispersată, dispersia ei și abaterea medie pătratică sunt mari.

Cazul *b* ocupă, după cum se vede, o poziție intermediară între cazurile *a* și *c*.

Pentru a cunoaște în modul cel mai rapid întregul ansamblu al legilor normale și pentru a putea aplica aceste legi, vom pleca dela două proprietăți fundamentale ale lor; aceste proprietăți, pe care le vor formula mai jos, nu le putem demonstra aici, întrucât ar trebui în prealabil să definim cu precizie legile normale, ceeace ar cere ca cititorul să aibă cunoștiințe de matematici superioare.

Proprietatea 1. Dacă variabila x este distribuită după o lege normală, atunci

1) pentru orice constante $c>0$ și d, mărimea $cx+d$ este deasemenea distribuită după o lege normală și

2) reciproc, pentru orice lege normală se poate găsi o pereche și numai una singură de numere $c>0$ și d, astfel încât mărimea $cx+d$ să fie distribuită chiar după această lege.

Astfel, dacă variabila întâmplătoare x este distribuită după o lege normală, legile de repartiție la care sunt supuse mărimile $cx+d$ pentru orice valori ale constantelor $c>0$ și d sunt deasemenea normale.

Proprietatea 2. Dacă variabilele întâmplătoare x și y sunt mutual independente și distribuite după legi normale, suma lor $z=x+y$ este deasemenea distribuită după o lege normală.

Acceptând fără demonstrație aceste două proprietăți fundamentale, putem da acum o motivare riguroasă unui număr de alte proprietăți ale legilor normale, rezultând din cele de mai sus, care sunt deosebit de importante din punct de vedere practic.

I. *Pentru orice pereche de numere a și $q>0$ există o singură lege normală cu o valoare medie a și cu o abatere medie pătratică q.*

In adevăr, fie x o variabilă întâmplătoare distribuită după o lege normală cu valoarea medie $\bar{x}$ și cu abaterea medie pătratică Q_x. In baza proprietății 1, afirmația noastră va fi demonstrată dacă putem să arătăm că există o singură pereche de numere $c>0$ și d, astfel încât mărimea $cx+d$ să aibă valoarea medie a și abaterea medie pătratică q. Dacă tabela variabilei este

x_1	x_2	. . .	x_n
p_1	p_2	. . .	p_n

(1)

atunci mărimii $cx+d$ (unde $c>0$ și d sunt deocamdată constante arbitrare) îi va corespunde tabela:

cx_1+d	cx_2+d	...	cx_n+d
p_1	p_2	...	p_n

Evident

$$\sum_k x_k p_k=\bar{x}, \quad \sum_k (x_k-\bar{x})^2 p_k=Q_x^2 \ ^1$$

condițiile impuse se reduc la relațiile

$$\sum_k (cx_k+d)\, p_k=a; \quad \sum_k (cx_k+d-a)^2 p_k=q^2.$$

Prima dintre aceste relații dă

$$c\sum_k x_k p_k+d\sum_k p_k=a,$$

sau

$$c\bar{x}+d=a, \tag{1}$$

iar a doua

$$\sum_k (cx_k+d-c\bar{x}-d)^2 p_k=c^2\sum_k (x_k-\bar{x})^2 p_k=c^2 Q_x^2=q^2,$$

de unde (intrucât $c>0$)

$$c=\frac{q}{Q_x}, \tag{2}$$

și deci, din (1) rezultă

$$d=a-c\bar{x}=a-\frac{q\bar{x}}{Q_x}. \tag{3}$$

Astfel, cunoscând pe a și q, putem determina întotdeauna în mod univoc numerele c și d, cu ajutorul formulelor (2) și (3); mărimea $cx+d$ se supune unei legi normale cu valoarea medie a și abaterea medie pătratică q; cu aceasta, afirmația noastră este în întregime demonstrată.

Dacă nu ne limităm la legile normale, ci examinăm toate legile de repartiție posibile, cunoașterea valorii medii și a dispersiei sau a abaterii medii pătratice a unei variabile întâmplătoare ne dă prea puține date despre legea ei de repartiție, întrucât există foarte multe asemenea legi (și încă fundamental deosebite între

[1] Semnul $\sum_k$ este o prescurtare pentru $\sum_{k=1}^{n}$.

ele) care admit toate aceeași valoare medie și aceeași dispersie; în cazul general, cunoașterea valorii medii și a dispersiei caracterizează doar foarte aproximativ legea de repartiție a variabilei întâmplătoare date.

Altfel se prezintă lucrurile dacă ne limităm numai la considerarea legilor normale.

După cum am văzut, orice presupunere relativă la valoarea medie și la dispersia unei variabile întâmplătoare date este pe deoparte compatibilă cu condiția ca ea să se supună unei legi normale. Pe de altă parte — și acesta este lucrul principal — dacă avem motive să bănuim dinainte că variabila dată este distribuită după o lege normală, atunci cunoscând valoarea medie și dispersia ei putem determina această lege în mod univoc, astfel încât natura ei ca variabilă întâmplătoare devine precis cunoscută. In particular, cunoscând valoarea medie și dispersia unei asemenea variabile, putem calcula probabilitatea ca valoarea ei să aparțină unui anumit interval ales arbitrar.

II. *Raportul dintre abaterea mediană (probabilă) și abaterea medie pătratică este același pentru toate legile normale.*

Să considerăm două legi normale arbitrare și o variabilă întâmplătoare x supusă primei dintre aceste legi. In baza proprietății fundamentale 1, există două constante $c>0$ și d astfel încât mărimea $cx+d$ să fie distribuită după a doua din legile date. Să însemnăm respectiv prin Q_x și E_x abaterea medie pătratică și abaterea mediană (probabilă) a primei variabile și prin q și e aceleași abateri pentru cea de a doua variabilă. Conform definiției abaterii probabile avem:

$$P\{|(cx+d)-(c\bar{x}+a)|<e\}=\frac{1}{2},$$

sau

$$P\{c|x-\bar{x}|<e\}=\frac{1}{2},$$

sau, în sfârșit,

$$P\left(|x-\bar{x}|<\frac{e}{c}\right)=\frac{1}{2}.$$

De aici, tot din definiția abaterii probabile, rezultă că $\frac{e}{c}$ este abaterea probabilă a variabilei x, adică

$$\frac{e}{c}=E_x,$$

de unde

$$\frac{e}{E_x}=c;$$

adică, din relația (2)

$$\frac{e}{E_x}=\frac{q}{Q_x},$$

de unde

$$\frac{e}{q}=\frac{E_x}{Q_x},$$

adică raportul dintre abaterea probabilă și abaterea medie pătratică este același pentru ambele legi.

Intrucât, conform ipotezei, aceste legi erau două legi normale absolut arbitrare, afirmația noastră este demonstrată.

Raportul $\frac{e}{q}$ este, prin urmare, o constantă și îl vom nota prin litera λ; s'a calculat că

$$\lambda\approx 0{,}674$$

Așa dar, pentru orice lege normală avem

$$e=0{,}674\ q.$$

Pe baza acestei legături extrem de simple între numerele e și q, în cazul variabilelor distribuite după legi normale, este practic indiferent pe care dintre cele două caracteristice ale dispersiei o vom folosi; am văzut mai sus că, în general (adică dacă nu ne limităm la variabile distribuite după legi normale), abaterea medie pătratică posedă o serie întreagă de proprietăți simple care lipsesc altor caracteristice, ceeace face ca, în majoritatea cazurilor, atât teoreticienii cât și practicienii să aleagă ca măsură a dispersiei tocmai abaterile medii pătratice. Am văzut deasemenea că artileriștii folosesc totuși aproape exclusiv abaterile mediane, și acum este clar dece această tradiție nu poate aduce niciun prejudiciu; variabilele întâmplătoare întâlnite în știința și în practica artileriei sunt distribuite aproape totdeauna după legi normale, iar pentru astfel de variabile, în virtutea proporționalității mai sus arătate, alegerea uneia sau alteia dintre cele două caracteristice este indiferentă din punct de vedere practic.

III. Fie x și y două variabile întâmplătoare independente supuse unor legi normale și $z=x+y$. Atunci

$$E_z=\sqrt{E_x^2+E_y^2},$$

unde E_x, E_y, E_z sunt respectiv abaterile probabile ale variabilelor x, y, z.

O formulă analoagă este valabilă pentru abaterile medii pătratice, după cum am văzut în § 25; oricare ar fi legile de

repartiție ale variabilelor x și y. In cazul când acestea sunt legi normale, în virtutea proprietății fundamentale 2, variabila z este deasemenea distribuită după o lege normală; deaceea, în baza proprietății anterioare II,

$$E_x=\lambda Q_x, \quad E_y=\lambda Q_y, \quad E_z=\lambda Q_z,$$

deci,

$$E_z=\lambda\sqrt{Q_x^2+Q_y^2}=\sqrt{(\lambda Q_x)^2+(\lambda Q_y)^2}=\sqrt{E_x^2+E_y^2}.$$

Așa dar, în cazul legilor normale una dintre cele mai importante proprietăți ale abaterilor medii pătratice se extinde nemijlocit și la abaterile probabile (mediane).

§ 32. Rezolvarea problemelor

Vom numi *legea normală fundamentală* — legea având valoarea medie egală cu zero și dispersia egală cu unitatea. Dacă x este o variabilă întâmplătoare supusă legii normale fundamentale, vom scrie pentru simplicitate

$$P\{|x|<a\}=\Phi(a).$$

pentru orice număr pozitiv a. Astfel $\Phi(a)$ este probabilitatea ca variabila x supusă legii normale fundamentale să nu depășească în valoare absolută numărul a. Pentru funcția $\Phi(a)$ s'au întocmit tabele foarte precise, care dau valorile ei pentru diferite valori ale numărului a. O asemenea tabelă este un ajutor prețios pentru oricine are lucrări legate de calculul probabilităților. Această tabelă se anexează la orice carte consacrată calculului probabilităților: cititorul o poate găsi și la sfârșitul acestei cărți. Având la îndemână tabela valorilor funcției $\Phi(a)$, se pot face ușor și cu multă precizie calculele pentru orice variabile distribuite după legi normale. Vom arăta acum prin exemple cum se face acest lucru.

Problema I. Variabila întâmplătoare x este distribuită, după o lege normală cu valoarea medie $\bar{x}$ și abaterea medie pătratică Q_x. Să se găsească probabilitatea ca abaterea $x-\bar{x}$ să nu depășească în valoare absolută numărul a.

Fie z o variabilă întâmplătoare distribuită după legea normală fundamentală. In virtutea proprietății fundamentale (1), se pot găsi două numere $c>0$ și d astfel încât mărimea $cz+d$ să aibă valoarea medie $\bar{x}$ și abaterea medie pătratică Q_x, adică să fie supusă aceleiași legi normale ca și variabila dată x. Deaceea

$$P(|x-\bar{x}|<a)=P[|(cz+d)-(c\bar{z}+d)|\leq a]=P(c|z-\bar{z}|<a);$$

dar, în virtutea formulei (2) dela pag. 98, $c=\frac{Q_x}{Q_z}=Q_x$, deoarece $Q_z=1$ (pentru legea normală fundamentală dispersia este egală cu 1). Obținem astfel

$$P(|x-\bar{x}|<a)=P(Q_x|z-\bar{z}|<a)=P\left(|z|<\frac{a}{Q_x}\right)=\Phi\left(\frac{a}{Q_x}\right). \quad (4)$$

Cu aceasta, problema este rezolvată, întrucât mărimea $\Phi\left(\frac{a}{Q_x}\right)$ poate fi găsită direct din tabelă. Astfel tabela ne permite să calculăm ușor cu ajutorul formulei (4) probabilitatea oricărei limite de abatere pentru o variabilă supusă unei legi normale.

Exemplul 1. La o mașină-unealtă se fabrică o piesă oarecare a cărei lungime x reprezintă o variabilă întâmplătoare distribuită după legea normală, având valoarea medie 20 cm și abaterea medie pătratică 0,2 cm. Să se calculeze probabilitatea ca lungimea piesei să fie cuprinsă între 19,7 cm și 20,3 cm, adică probabilitatea ca abaterile într'un sens sau altul să nu depășească 0,3 cm.

In conformitate cu formula (4) și cu tabela noastră

$$P\{|x-20|<0{,}3\}=\Phi\left(\frac{0{,}3}{0{,}2}\right)=\Phi(1{,}5)=0{,}866.$$

Astfel, circa 87% din toate produsele fabricate în condițiile date vor avea lungimea cuprinsă între 19,7 și 20,3 cm; restul de 13% vor avea abateri mai mari dela medie.

Exemplul 2. In condițiile exemplului 1 să se afle ce precizie se poate garanta cu probabilitatea 0,95 pentru lungimea produsului dat.

Problema revine, evident, la căutarea unui număr pozitiv a pentru care

$$P\{|x-20|<a\}>0{,}95.$$

Calculul din exemplul 1 arată că valoarea $a=0{,}3$ este prea mică pentru aceasta căci, în acest caz, membrul din stânga al inegalității de mai sus este mai mic decât 0,87. Intrucât, conform ecuației (4)

$$P\{|x-20|<a\}=\Phi\left(\frac{a}{0{,}2}\right)=\Phi(5a),$$

trebue să găsim mai întâi în tabela noastră o valoare $5a$ astfel încât

$$\Phi(5a)>0{,}95.$$

Scanned with OKEN Scanner

Se găsește că aceasta are loc pentru

$$5a > 1{,}97,$$

de unde $a > 0{,}394$.

Astfel, cu o probabilitate care depășește 0,95, putem garanta că abaterea în lungime nu va depăși 0,4 cm.

Exemplul 3. In unele probleme practice se consideră că variabila întâmplătoare x distribuită după o lege normală are abateri mai mici decât trei abateri medii pătratice Q_x. Este întemeiată această afirmație?

Formula (4) și tabela noastră ne arată că

$$P\{|x-\bar{x}|<3Q_x\}=\Phi(3)>0{,}997,$$

și deci

$$P\{|x-\bar{x}|>3Q_x\}<0{,}003.$$

Din punct de vedere practic, aceasta înseamnă că abaterile care depășesc în valoare absolută $3Q_x$ se vor întâlni mai rar decât trei la mie. Dacă putem sau nu neglija o astfel de posibilitate depinde, firește, de condițiile problemei și aceasta nu se poate stabili odată pentru totdeauna.

Observăm că relația

$$P\{|x-\bar{x}|<3Q_x\}=\Phi \qquad (3)$$

este, după cum se vede, un caz particular al formulei

$$P\{|x-\bar{x}|<aQ_x\}=\Phi(a) \qquad (5)$$

care rezultă din formula (4) și are loc pentru orice variabilă întâmplătoare x distribuită după o lege normală.

Exemplul 4. La o greutate medie a proectilului de 8,4 kg s'a constatat că abaterile depășind în valoare absolută 50 g se întâlnesc în medie de trei ori la fiecare 100 proectile. Presupunând că greutatea proectilului este distribuită după o lege normală, să se determine abaterea ei probabilă.

Din datele problemei știm că:

$$P(|x-8{,}4|>0{,}05)=0{,}03,$$

unde x este greutatea unui proectil ales la întâmplare. De aici

$$0{,}97=P(|x-8{,}4|<0{,}05)=\Phi\left(\frac{0{,}05}{Q_x}\right);$$

tabela ne arată că $\Phi(a)=0{,}97$ pentru $a\approx 2{,}17$.

Scanned with OKEN Scanner

Deaccea

$$\frac{0,05}{Q_x}=2,17,$$

de unde

$$Q_x\approx\frac{0,05}{2,17}.$$

După cum știm abaterea probabilă este

$$E_x=0,674\,Q_x\approx 0,0155\ \text{kg}=15,5\ \text{g}.$$

Exemplul 5. La o tragere de artilerie, abaterea proectilului dela țintă este determinată de trei cauze mutual independente: 1) eroarea în determinarea poziției țintei; 2) eroarea de ochire; 3) eroarea datorită unor cauze care se schimbă dela un foc la altul (greutatea proectilului, condițiile atmosferice, etc.). Presupunând că toate cele trei erori sunt distribuite după legi normale, cu valoarea medie 0 și că abaterile lor probabile sunt respectiv egale cu 24 m, 8 m și 12 m, să se afle probabilitatea ca abaterea totală dela țintă să nu depășească 40 m.

Intrucât în virtutea proprietății III (pag. 100) abaterea probabilă a erorii totale x este

$$\sqrt{24^2+8^2+12^2}=28\ \text{m},$$

abaterea medie pătratică a erorii totale este

$$\frac{28}{0,674}\approx 41,5$$

deci

$$P(|x|<40)=\Phi\left(\frac{40}{41,5}\right)\approx\Phi\,(0,964)=0,665.$$

Abateri care nu depășesc 40 m se vor observa deci cam în $\frac{2}{3}$ din totalitatea cazurilor.

Problema II. Variabila întâmplătoare x este distribuită după o lege normală cu valoarea medie $\bar{x}$ și abaterea medie pătratică Q_x. Să se afle probabilitatea ca valoarea absolută a abaterii $x-\bar{x}$ să fie cuprinsă între numerele a și b $(0<a<b)$.

Intrucât, conform regulii adunării, avem

$$P(|x-\bar{x}|<b)=P(|x-\bar{x}|<a)+P(a<|x-\bar{x}|<b),$$

rezultă

$$P(a<|x-\bar{x}|<b)=P(|x-\bar{x}|<b)-P(|x-\bar{x}|<a)=$$
$$=\Phi\left(\frac{b}{Q_x}\right)-\Phi\left(\frac{a}{Q_x}\right), \qquad (6)$$

ceeace rezolvă problema pusă.

Scanned with OKEN Scanner

Pentru covârșitoarea majoritate a cerințelor practicii, tabela valorilor funcției $\Phi(a)$ de care ne-am servit tot timpul este totuși un instrument greoi de calcul. Deobicei este necesar să se calculeze numai probabilitățile ca abaterea $x-\bar{x}$ să fie cuprinsă în intervale mai mult sau mai puțin lungi: deaceea, în practică este de dorit ca pe lângă tabela „completă" să avem și tabele prescurtate, care sunt ușor de întocmit din tabela completă cu ajutorul formulei (6).

Iată un exemplu de modul cum se construește o asemenea tabelă, mult mai primitivă decât tabela dela sfârșitul cărții, dar totuși satisfăcătoare în multe cazuri. Să împărțim întregul interval de variație a mărimii $x-\bar{x}$ în 5 părți: 1) dela zero până la $0{,}32\ Q_x$; 2) dela $0{,}32\ Q_x$ până la $0{,}69\ Q_x$; 3) dela $0{,}69\ Q_x$ până la $1{,}15\ Q_x$; 4) dela $1{,}15\ Q_x$ până la $2{,}58\ Q_x$ și 5) peste $2{,}58\ Q_x$.

Aplicând formula (4) obținem:

$$P(|x-\bar{x}|<0{,}32\,Q_x)=\Phi(0{,}32)\approx 0{,}25;$$

$$P(0{,}32\,Q_x<|x-\bar{x}|<0{,}69\,Q_x)=\Phi(0{,}69)-\Phi(0{,}32)\approx 0{,}25;$$

$$P(0{,}69\,Q_x<|x-\bar{x}|<1{,}15\,Q_x)=\Phi(1{,}15)-\Phi(0{,}69)\approx 0{,}25;$$

$$P(1{,}15\,Q_x<|x-\bar{x}|<2{,}58\,Q_x)=\Phi(2{,}58)-\Phi(1{,}15)\approx 0{,}24;$$

$$P(|x-\bar{x}|>2{,}58\,Q_x)=1-\Phi(2{,}58)\approx 0{,}01.$$

Rezultatul acestui calcul poate fi reprezentat comod cu ajutorul unei scheme grafice (fig. 16).

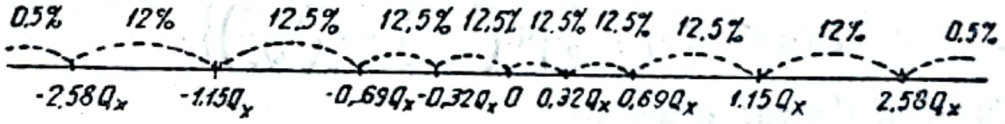

Fig. 16.

Intreaga dreaptă nelimitată este aici împărțită în 10 intervale, 5 pozitive și 5 negative. Deasupra fiecărui interval se indică procentul abaterilor reale observate, care vor fi cuprinse în medie în el. Astfel, conform calculelor făcute mai sus, în intervalele $(1{,}15\ Q_x, -0{,}69\ Q_x)$ și $(0{,}69\ Q_x, 1{,}15\ Q_x)$ luate laolaltă, trebue să se afle circa 25% din totalul abaterilor. Din cauza simetriei legilor normale, abaterile vor cădea aproximativ la fel de frecvent în ambele intervale, astfel încât în fiecare dintre ele vor cădea circa 12,5% din numărul total al abaterilor. Cu ajutorul acestei scheme simple sau al alteia asemănătoare, ne putem face o ideie directă asupra trăsăturilor esențiale ale repartiției abaterilor unei variabile întâmplătoare supusă unei legi normale, cu o valoare medie și cu o abatere medie pătratică arbitrare.

Să vedem acum, în sfârșit, cum se calculează probabilitatea ca o variabilă întâmplătoare supusă unei legi normale să ia valori cuprinse într'un anumit interval dinainte stabilit.

Problema III. Știind că variabila întâmplătoare x este distribuită după o lege normală (cu valoarea medie $\bar{x}$ și abaterea medie pătratică Q_x), să se calculeze cu ajutorul tabelelor probabilitatea inegalității $a<x<b$, unde a și b $(a<b)$ sunt două numere arbitrare date.

Va trebui să examinăm trei cazuri în funcție de poziția numerelor a și b față de $\bar{x}$.

Cazul întâi: $\bar{x} \leq a \leq b$.

Conform regulii de adunare

$$P(\bar{x}<x<b) = P(\bar{x}<x<a) + P(a<x<b)$$

de unde

$$P(a<x<b) = P(\bar{x}<x<b) - P(\bar{x}<x<a) =$$
$$= P(0<x-\bar{x}<b-\bar{x}) - P(0<x-\bar{x}<a-\bar{x}).$$

Dar în virtutea simetriei legilor normale, pentru un $\alpha > 0$ oarecare obținem:

$$P(0<x-\bar{x}<\alpha) = P(-\alpha<x-\bar{x}<0) =$$
$$= \frac{1}{2} P(-\alpha<x-\bar{x}<\alpha) = \frac{1}{2} P(|x-\bar{x}|<\alpha) = \frac{1}{2}\Phi\left(\frac{\alpha}{Q_x}\right); \qquad (7)$$

deci

$$P(a<x<b) = \frac{1}{2}\left\{\Phi\left(\frac{b-\bar{x}}{Q_x}\right) - \Phi\left(\frac{a-\bar{x}}{Q_x}\right)\right\}.$$

Cazul al doilea: $a \leq \bar{x} \leq b$.

$$P(a<x<b) = P(a<x<\bar{x}) + P(\bar{x}<x<b) = P(a-\bar{x}<x-\bar{x}<0) +$$
$$+ P(0<x-\bar{x}<b-\bar{x}) = \frac{1}{2}\left\{\Phi\left(\frac{\bar{x}-a}{Q_x}\right) + \Phi\left(\frac{b-\bar{x}}{Q_x}\right)\right\},$$

în virtutea formulei (7).

Cazul al treilea: $a \leq b \leq \bar{x}$,

$$P(a<x<\bar{x}) = P(a<x<b) + P(b<x<\bar{x}),$$

de unde

$$P(a<x<b) = P(a<x<\bar{x}) - P(b<x<\bar{x}) = P(a-\bar{x}<x-\bar{x}<0) -$$
$$- P(b-\bar{x}<x-\bar{x}<0) = \frac{1}{2}\left\{\Phi\left(\frac{\bar{x}-a}{Q_x}\right) - \Phi\left(\frac{\bar{x}-b}{Q_x}\right)\right\}.$$

Problema este astfel rezolvată în toate cele trei cazuri posibile. După cum se vede, pentru o variabilă întâmplătoare distribuită după o lege normală oarecare, tabela noastră ne permite să cal-

Scanned with OKEN Scanner

culăm probabilitatea ca variabila să ia valori în orice interval dat şi prin aceasta caracterizează pe deplin legea ei de repartiţie.

Pentru a vedea cum se fac calculele în practică, să examinăm următorul exemplu.

Exemplu. Se execută o tragere de artilerie din punctul O în lungul dreptei OX. Distanţa medie străbătută de proectil este de 1 200 m. Presupunând că această distanţă H este distribuită după o lege normală cu abaterea medie pătratică de 40 m, să se determine procentul proectilelor trase care dă o depăşire cuprinsă între 60 şi 80 m.

Pentru ca obuzul să aibă o asemenea depăşire, trebue să avem $1\,260<H<1\,280$; aplicând ultima formulă din cazul întâi al problemei III, obţinem

$$P(1\,260<H<1\,280)=\frac{1}{2}\left\{\Phi\left(\frac{1\,280-1\,200}{40}\right)-\Phi\left(\frac{1\,260-1\,200}{40}\right)\right.=$$

$$=\frac{1}{2}\left\{\Phi(2)-\Phi(1{,}5)\right\};$$

tabela ne dă

$$\Phi(2)\approx 0{,}955, \quad \Phi(1{,}5)\approx 0{,}866,$$

de unde

$$P(1\,260<H<1\,280)\approx 0{,}044;$$

constatăm aşa dar că ceva mai mult de 4% din proectilele trase vor cădea cu depăşirea cerută.

INCHEIERE

Calculul probabilităţilor este una din ramurile matematicii, care s'a desvoltat cu cea mai mare rapiditate în ultimele decenii. Rezultatele teoretice găsite în ultimul timp deschid noi posibilităţi pentru utilizarea metodelor calculului probabilităţilor în ştiinţele naturii. O cercetare mai precisă, mai amănunţită a fenomenelor naturii conduce în acelaşi timp calculul probabilităţilor la descoperirea unor metode noi, unor legi noi. Calculul probabilităţilor este o ramură a matematicii care nu este ruptă de viaţă şi nici de cerinţele altor ştiinţe, ci merge în pas cu desvoltarea generală a ştiinţelor naturii şi a tehnicei. Cititorul nu trebue să înţeleagă greşit cele spuse mai sus şi să creadă că teoria probabilităţilor s'a transformat acum într'un sprijin, într'un mijloc auxiliar pentru rezolvarea problemelor practice. Nu este deloc aşa: în ultimele trei decenii, calculul probabilităţilor a devenit o ştiinţă matematică închegată, cu problemele ei proprii şi cu metodele ei proprii de cercetare. Totodată, cele mai importante şi naturale probleme ale calculului probabilităţilor s'au dovedit utile pentru rezolvarea unor probleme actuale ale ştiinţelor naturii.

Marea desvoltare a calculului probabilităţilor este strâns legată de tradiţiile şi de realizările ştiinţei ruse. In timp ce în Apus, în secolul trecut, teoria probabilităţilor intrase într'un impas, în Rusia genialul matematician P. L. Cebâşev a găsit o cale nouă pentru desvoltarea ei mai departe şi anume, studierea multilaterală a şirurilor variabile întâmplătoare independente. Cebâşev şi elevii săi A. A. Marcov şi A. M. Liapunov au obţinut pe această cale rezultate fundamentale (legea numerelor mari, teorema lui Liapunov).

Cititorii, au luat cunoştinţă, în cele arătate până acum de legea numerelor mari; vom căuta acum să prezentăm în linii mari o altă propoziţie importantă din calculul probabilităţilor, numită teorema lui Liapunov (sau teorema centrală a limitei din calculul probabilităţilor).

Scanned with OKEN Scanner

Marea ei importanță constă în faptul că numeroase fenomene, al căror rezultat depinde de întâmplare, au loc în esență după următoarea schemă: fenomenul studiat se află sub acțiunea unui număr foarte mare de cauze întâmplătoare, acționând independent, fiecare dintre ele având o influență foarte mică asupra evenimentului în ansamblul lui. Acțiunea fiecăreia dintre aceste cauze este exprimată de variabilele întâmplătoare $\xi_1, \xi_2, \ldots \xi_n$, iar influența lor globală asupra fenomenului este egală cu suma $s_n = \xi_1 + \xi_2 + \ldots + \xi_n$. Întrucât calcularea influenței fiecăreia dintre aceste cauze (cu alte cuvinte darea funcțiilor de repartiție ale variabilelor ξ) și chiar simpla lor enumerare este practic imposibilă; rezultă în mod clar importanța elaborării unor metode care să permită studierea acțiunii lor globale, indiferent de natura fiecărui factor în parte. Metodele obișnuite de cercetare, pentru soluționarea acestei probleme sunt ineficace, ele trebue să fie înlocuite cu alte metode, pentru care numărul mare de cauze care influențează fenomenul să nu fie o piedică, ci dimpotrivă să ușureze rezolvarea problemei puse. Aceste metode trebue să compenseze insuficienta cunoaștere a fiecărei cauze în parte, prin numărul lor mare, prin masivitatea lor. Teorema limită centrală stabilită în lucrările academicianului P. L. Cebâșev (1821—1894), A. A. Marcov (1856 până 1922) și A. M. Liapunov (1857—1918) afirmă că, dacă toate cauzele ce acționează $\xi_1, \xi_2, \ldots \xi_n$ sunt mutual independente, dacă numărul lor n este foarte mare și dacă acțiunea fiecăreia dintre aceste cauze față de acțiunea lor globală este neînsemnată, legea de repartiție a sumei s_n poate diferi, oricât de puțin, de legea de repartiție.

Să dăm exemple de fenomene ce au loc după schema descrisă mai sus.

La o tragere de artilerie într'o țintă este inevitabil ca punctul de nimerire al proectilului să prezinte abateri dela țintă. Acesta este fenomenul bine cunoscut al dispersiei obuzelor. Întrucât dispersia este rezultatul acțiunii unui număr foarte mare de factori independenți (de exemplu iregularități în strunjirea corpului și vârfului proectilului, oscilații în densitatea materialului din care este strunjit vârful proectilului, fluctuații neînsemnate ale cantității de explozibil în diferitele proectile, erori imperceptibile la ochire, schimbări foarte mici în starea atmosferică la diferitele trageri și multe alte cauze), fiecare dintre ei având o influență extrem de mică asupra traectoriei proectilului, rezultă din teorema lui Liapunov că ea se va supune legii normale. De acest rezultat se ține seamă în teoria tragerilor și el stă la baza elaborării regulilor de tragere.

Scanned with OKEN Scanner

Atunci când facem o observație cu scopul de a măsura o anumită constantă fizică, rezultatele observației noastre sunt inevitabil influențate de un număr foarte mare de cauze, care nu pot fi evaluate separat, dar care dau naștere la erori de măsură. Dintre aceste erori fac parte cele provenite din starea aparatului de măsură, ale cărui indicații pot varia puțin (insensibil) sub influența diferitelor cauze (atmosferice, termice, mecanice, etc.). Dintre acestea mai fac parte erorile observatorului, provocate de particularități ale văzului sau auzului său și care, deasemenea, se schimbă pe nesimțite în funcție de starea psihică sau fizică a observatorului. Eroarea efectivă a măsurării este astfel o eroare rezultantă a unui număr foarte mare de erori neinsemnate ca mărime, independente între ele sau, cum se spune, elementare, depinzând de greșeli întâmplătoare. In virtutea teoremei lui Liapunov putem din nou să ne așteptăm ca erorile de observație să urmeze legile de repartiție normale.

Putem cita oricâte exemple asemănătoare: poziția și viteza unei molecule de gaz, determinate de un mare număr de ciocniri cu alte molecule; cantitatea de substanță care a difuzat; abaterea pieselor unei mașini dela dimensiunile stabilite într'o producție de masă a mașinilor; distribuția creșterii animalelor, a plantelor sau a altor organisme, etc.

Perfecționarea fizicei statistice și a unui număr de ramuri ale tehnicei a pus în fața calculului probabilităților numeroase probleme cu totul noi care nu intră în cadrul schemelor clasice. In timp ce fizica sau tehnica își îndreaptă atenția către studiul *procesului,* adică al fenomenului care se petrece în timp, calculul probabilităților nu a avut elaborate nici procedee generale nici scheme particulare pentru rezolvarea problemelor ce se ivesc la studiul unor astfel de fenomene. A apărut astfel necesitatea imperioasă de a elabora o teorie generală a *proceselor întâmplătoare,* adică o teorie care să studieze variabilele întâmplătoare, *care depind de unul sau mai mulți parametri continui.*

Să enumerăm câteva probleme care duc la studierea variabilelor întâmplătoare care variază în timp. Să presupunem că vrem să urmărim mișcarea unei molecule oarecare de gaz sau lichid. Această moleculă la momente întâmplătoare se ciocnește de alte molecule, schimbându-și cu acest prilej viteza și direcția mișcării. Această modificare a stării moleculei este supusă la fiecare moment unei acțiuni întâmplătoare. O serie întreagă de fenomene fizice cer ca pentru studiul lor să se poată calcula probabilitatea ca un anumit număr de molecule într'un interval de timp oarecare să se poată deplasa cu o distanță anumită. Astfel, dacă au fost aduse în contact două gaze sau două lichide, are loc pătrunderea

reciprocă a moleculelor unui lichid în celălalt: se petrece difuziunea. Cât timp durează procesul difuziunii? După care legi? Când amestecul de gaze devine practic omogen? La toate aceste întrebări teoria statistică a difuziunii ne dă răspunsuri la baza cărora stă calculul probabilităţilor, studierea proceselor întâmplătoare.

Evident că aceeaşi problemă se pune şi în chimie când se studiază procesul interacţiunii chimice a substanţelor, procesul unei reacţii chimice. Care este partea de molecule care a intrat deja în reacţie, cum decurge reacţia în timp, când este terminată reacţia din punct de vedere practic?

Un mare număr de fenomene fizice decurg după principiul desintegrării radioactive. Acest fenomen constă în aceea că atomii substanţei radioactive se desintegrează devenind atomii unui alt element. Fiecare desintegrare de atomi are loc instantaneu, ca o explozie, cu eliberarea unei anumite cantităţi de energie. Numeroase observaţii au arătat că desintegrarea atomilor are loc la momente întâmplătoare şi independente unul de altul (cu condiţia ca masa substanţei radioactive să nu fie prea mare). Pentru studierea procesului desintegrării radioactive, este foarte important să se determine probabilitatea ca într'un interval oarecare de timp să se descompună un anumit număr de atomi. Aceasta este o problemă tipică a teoriei proceselor întâmplătoare. Din punct de vedere formal, dacă am căuta numai explicarea imaginii matematice a fenomenelor, alte fenomene se petrec în mod analog: gradul de încărcare a centralei telefonice, adică numărul de apeluri sosite la centrala telefonică dela abonaţi, mişcarea browniană, ruperea firelor la un război de ţesut, etc.

Bazele teoriei generale a proceselor întâmplătoare au fost puse de lucrările fundamentale al matematicienilor sovietici A. N. Colmogorov şi A. I. Hincin, în primii ani după 1930. Ceva mai devreme, în primul deceniu al secolului nostru, A. A. Marcov a început să studieze şirurile de variabile întâmplătoare dependente, care au căpătat denumirea de lanţuri Marcov. Teoria elaborată de el doar ca o disciplină matematică s'a transformat în deceniul al treilea, în mâinile fizicienilor, într'un instrument activ pentru studierea naturii. De atunci mulţi savanţi (S. N. Bernştein, V. I. Romanovschi, A. N. Colmogorov, J. Hadamard, M. Frechet, etc.) au adus contribuţii importante la teoria lanţurilor Marcov.

In deceniul al treilea, A. N. Colmogorov, E. E. Sluţchi, A. I. Hincin şi Paul Lévy au găsit legătura strânsă dintre teoria probabilităţilor şi disciplinele matematice care studiază mulţimile şi noţiunea generală de funcţie (teoria mulţimilor şi teoria funcţiilor de variabilă reală).

Descoperirea acestei legături s'a dovedit extrem de fecundă și tocmai pe această cale s'au putut rezolva definitiv problemele clasice puse de Cebâșev.

In sfârșit, mai trebue să menționăm tot aici lucrările lui S. N. Bernștein și A. N. Colmogorov, de sistematizare a noțiunilor fundamentale ale calculului probabilităților.

In încheiere, putem spune că în progresul general al calculului probabilităților rolul hotărîtor l-a jucat știința rusă.

Putem spune cu mândrie că ideile fundamentale ale calculului probabilităților sau au apărut la noi în țară, sau au căpătat în țara noastră o puternică desvoltare ori strălucita lor desăvârșire.

Scanned with OKEN Scanner

Tabela valorilor funcției $\Phi(a)$

a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$
0,00	0,000	**0,40**	0,311	**0,80**	0,576	**1,20**	0,770	**1,60**	0,890
0,01	0,008	0,41	0,318	0,81	0,582	1,21	0,774	1,61	0,893
0,02	0,016	0,42	0,326	0,82	0,588	1,22	0,778	1,62	0,895
0,03	0,024	0,43	0,333	0,83	0,593	1,23	0,781	1,63	0,897
0,04	0,032	0,44	0,340	0,84	0,599	1,24	0,785	1,64	0,899
0,05	0,040	0,45	0,347	0,85	0,605	1,25	0,789	1,65	0,901
0,06	0,048	0,46	0,354	0,86	0,610	1,26	0,792	1,66	0,903
0,07	0,056	0,47	0,362	0,87	0,616	1,27	0,796	1,67	0,905
0,08	0,064	0,48	0,369	0,88	0,621	1,28	0,800	1,68	0,907
0,09	0,072	0,49	0,377	0,89	0,627	1,29	0,803	1,69	0,909
0,10	0,080	**0,30**	0,383	**0,90**	0,632	**1,30**	0,806	**1,70**	0,911
0,11	0,088	0,51	0,390	0,91	0,637	1,31	0,810	1,71	0,913
0,12	0,096	0,52	0,397	0,92	0,642	1,32	0,813	1,72	0,915
0,13	0,103	0,53	0,404	0,93	0,648	1,33	0,816	1,73	0,916
0,15	0,111	0,54	0,411	0,94	0,653	1 34	0,820	1,74	0,918
0,15	0,119	0,55	0,418	0,95	0,658	1,35	0,823	1,75	0,920
0,16	0,127	0,56	0,425	0,96	0,663	1,36	0,826	1,76	0,922
0,17	0,135	0,57	0,431	0,97	0,668	1,37	0,829	1,77	0,923
0,18	0,143	0,58	0,438	0,98	0,673	1,38	0,832	1,78	0,925
0,19	0,151	0,59	0,445	0,99	0,678	1,39	0,835	1,79	0,927
0,20	0,159	**0,60**	0,451	**1,00**	0,683	**1,40**	0,838	**1,80**	0,928
0,21	0,166	0,61	0,458	1,01	0,688	1,41	0,841	1,81	0,930
0,22	0,174	0,62	0,465	1,02	0,692	1,42	0,844	1,82	0,931
0,23	0,182	0,63	0,471	1,03	0,697	1,43	0,847	1,83	0,933
0,24	0,190	0,64	0,478	1,04	0,702	1,44	0,850	1,84	0,934
0,25	0,197	0,65	0,484	1,05	0,706	1,45	0,853	1,85	0,936
0,26	0,205	0,66	0,491	1,06	0,711	1,46	0,856	1,86	0,937
0,27	0,213	0,67	0,497	1,07	0,715	1,47	0,858	1,87	0,939
0,28	0,221	0,68	0,504	1,08	0,720	1,48	0,861	1,88	0,940
0,29	0,228	0,69	0,510	1,09	0,724	1,49	0,864	1,89	0,941
0,30	0,236	**0,70**	0,516	**1,10**	0,729	**1,50**	0,866	**1,90**	0,943
0,31	0,243	0,71	0,522	1,11	0,733	1,51	0,867	1,91	0,944
0,32	0,251	0,72	0,528	1,12	0,737	1,52	0,871	1,92	0,945
0,33	0,259	0,73	0,535	1,13	0,742	1,53	0,874	1,93	0,946
0,34	0,266	0,74	0,541	1,14	0,746	1,54	0,876	1,94	0,948
0,35	0,274	0,75	0,547	1,15	0,750	1,55	0,879	1,95	0,949
0,36	0,281	0,76	0,553	1,16	0,754	1,56	0,881	1,96	0,950
0,37	0,289	0,77	0,559	1,17	0,758	1,57	0,884	1,97	0,951
0,38	0,296	0,78	0,565	1,18	0,762	1,58	0,886	1,98	0,952
0,39	0,303	0,79	0,570	1,19	0,766	1,59	0,888	1,99	0,953

Scanned with OKEN Scanner

a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$	a	$\Phi(a)$
2,00	0,955	**1,20**	0,972	**2,40**	0,984	**2,60**	0,991	2,88	0,996
2,01	0,956	2,21	0,973	2,41	0,984	2,61	0,991	**2,90**	0,996
2,02	0,957	2,22	0,974	2,42	0,984	2,62	0,991	2,92	0,996
2,03	0,958	2,23	0,974	2,43	0,985	2,63	0,991	2,94	0,997
2,04	0,959	2,24	0,975	2,44	0,985	2,64	0,992	2,96	0,997
2,05	0,960	2,25	0,076	2,45	0,986	2,65	0,992	2,98	0,997
2,06	0,961	2,26	0,976	2,46	0,986	2,66	0,992		
2,07	0,962	2,27	0,977	2,47	0,986	2,67	0,992	**3,00**	0,997
2,08	0,962	2,28	0,977	2,48	0,987	2,68	0,993	3,10	0,998
2,09	0,963	2,29	0,978	2,49	0,987	2,69	0,993	3,20	0,999
								3,30	0,999
2,10	0,964	**2,30**	0,979	**2,50**	0,988	**2,70**	0,993	3,40	0,999
2,11	0,965	2,31	0,979	2,51	0,988	2,72	0,993		
2,12	0,966	2,32	0,980	2,52	0,988	2,74	0,994	**3,50**	0,9995
2,13	0,967	2,33	0,980	2,53	0,989	2,76	0,994	3,60	0,9997
2,14	0,968	2,34	0,981	2,54	0,989	2,78	0,995	3,70	0,9998
2,15	0,968	2,35	0,981	2,55	0,989			3,80	0,99986
2,16	0,969	**2,36**	0,982	2,56	0,990	**2,80**	0,995	3,90	0,99990
2,17	0,970	2,37	0,982	2,57	0,990	2,82	0,995	**4,00**	0,99994
2,18	0,971	2,38	0,983	2,58	0,990	2,84	0,995		
2,19	0,971	2,39	0,983	2,59	0,990	2,86	0,996	**5,00**	0,99999994

Scanned with OKEN Scanner

TABLA DE MATERII

Pag.

Prefața la prima ediție 3
Prefețele la a doua și a treia ediție 4

Partea întâia

PROBABILITĂȚI

Cap. I. **Probabilitățile evenimentelor** 5
§ 1. Noțiunea de probabilitate 5
§ 2. Evenimente imposibile și evenimente sigure 9
§ 3. Problemă 10

Cap. II. **Regula de adunare a probabilităților** 11
§ 4. Deducerea regulii de adunare a probabilităților 11
§ 5. Sistem complect de evenimente 14
§ 6. Exemple 16

Cap. III. **Probabilități condiționate și regula de înmulțire** 18
§ 7. Noțiunea de probabilitate condiționată 18
§ 8. Deducerea regulii de înmulțire a probabilităților 20
§ 9. Evenimente independente 22

Cap. IV. **Consecințe ale regulilor de adunare și înmulțire** 26
§ 10. Stabilirea câtorva inegalități 26
§ 11. Formula probabilității totale 28
§ 12. Formula lui Bayes.. 31

Cap. V. **Schema lui Bernoulli** 36
§ 13. Exemple 36
§ 14. Formulele lui Bernoulli 38
§ 15. Numărul cel mai probabil de apariții ale unui eveniment 41

Cap. VI. **Teorema lui Bernoulli** 47
§ 16. Conținutul teoremei lui Bernoulli 47
§ 17. Demonstrația teoremei lui Bernoulli 48

Partea a doua

VARIABILE ÎNTÂMPLĂTOARE

Cap. VII. **Variabile întâmplătoare și legi de repartiție** 54
§ 18. Noțiunea de variabilă întâmplătoare 54
§ 19. Noțiunea de lege de repartiție 56

Pag.

Cap. VIII. **Valori medii** 59
§ 20. Definiţia valorii medii a unei variabile întâmplătoare 59

Cap. IX. **Valorile medii ale sumei şi ale produsului** 68
§ 21. Teorema valorii medii a sumei 68
§ 22. Teorema valorii medii a produsului 71

Cap. X. **Dispersie şi abateri medii** 73
§ 23. Insuficienţa valorii medii pentru caracterizarea unei variabile întâmplătoare 73
§ 24. Diferite metode de măsurare a dispersiei unei variabile întâmplătoare.. 74
§ 25. Teoreme despre abaterea medie pătratică 80

Cap. XI. **Legea numerelor mari** 85
§ 26. Inegalitatea lui Cebâşev 85
§ 27. Legea numerelor mari 86
§ 28. Demonstraţia legii numerelor mari 88

Cap. XII. **Legi normale** 91
§ 29. Modul de a pune problema 91
§ 30. Noţiunea de curbă de repartiţie 93
§ 31. Proprietăţile curbelor normale de repartiţie 95
§ 32. Rezolvarea problemelor 101

Incheiere 103

Anexa. Tabela valorilor funcţiei $\Phi(u)$ 113

Redactor de carte: Neicov Valentina
Telnoredactor: Kivu Niculae
Corector resp.: Ing. Rotaru Vasile

Dat la cules 14. VII. 1953. Bun de tipar 21. X. 1953. Hârtia velină 61×86/16—65 g./m². Coli editoriale 7,25. Coli tipar 7,25. Preţul unui exemplar broşat lei 4,07. C-da Nr. 2166—A. 03162/1953. Pentru bibliotecile mici, indicele de clasificare 519. Pentru bibliotecile mari, indicele de clasificare 519.2.

Tiparul executat la
Intreprinderea Poligrafică Sibiu, Str. I.V. Stalin 15 - RPR

Scanned with OKEN Scanner

ERATĂ

Pag.	Rândul	In loc de	Se va citi	Din vina
27	2 de sus	$(\overline{A}$ şi $\overline{B})$	$(A$ şi $\overline{B})$	Redacţiei
29	11 de jos	$P(\overline{B}) \cdot P_B(A)$	$P(\overline{B}) \cdot P_{\overline{B}}(A)$	Tipografiei
40	6 de sus	C^k	C_n^k	Tipografiei
43	8 de jos	$k \geq$	$k_0 \geq$	Redacţiei
48	2 de jos	n	εn	Tipografiei
78	7 de jos	$Q_x^2 \geq M_x$	$Q_x^2 \geq M_x^2$	Tipografiei
83	10 de sus	$\sum_{k=1}^{n} a\,k$	$\sum_{k=1}^{n} a_k$	Redacţiei
85	4 de jos	nu permite	ne permite	Tipografiei
99	10 de jos	$-(c\overline{x}+a)$	$-(c\overline{x}+d)$	Tipografiei
102	4 de sus	$P(Qx \mid z-\overline{z} \mid <a)$	$P(Q_x \mid z-\overline{z} \mid <a)$	Redacţiei
103	15 de jos	$=\Phi$	$=\Phi(3)$	Redacţiei
109	18 de jos	legea de	legea normală de	Redacţiei

Scanned with OKEN Scanner